

GenCore version 5.1.1.8  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2006, 10:55:27 ; Search time 201 Seconds  
(without alignments)  
646.017 Million cell updates/sec

Title: US-10-768-158-2  
Perfect score: 1530  
Sequence: 1 MAESEAETPSTPGFEFSKYF.....FDLVYKQMGKDLTFDFYL 284

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

- Database : A\_Geneseq\_8.\*
- 1: Geneseqp1980s.\*
  - 2: Geneseqp1990s.\*
  - 3: Geneseqp2000s.\*
  - 4: Geneseqp2001s.\*
  - 5: Geneseqp2002s.\*
  - 6: Geneseqp2003as.\*
  - 7: Geneseqp2003bs.\*
  - 8: Geneseqp2004s.\*
  - 9: Geneseqp2005s.\*
  - 10: Geneseqp2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1530	100.0	284	3	AAV79215 Human tra
2	1530	100.0	284	5	AAU78268 Human nar
3	1530	100.0	284	6	ABR01809 Human can
4	1530	100.0	284	8	ADQ89158 Human uro
5	1530	100.0	284	8	ADR44884 Polypepti
6	1530	100.0	284	9	ADZ21933 Human SUL
7	1506	98.4	284	6	ABM04807 Rat sulfo
8	1002	65.5	218	6	ABR01860 Human can
9	867.5	56.7	171	6	ABR01861 Human can
10	867.5	56.7	171	9	ADZ21934 Human SUL
11	850.5	55.6	159	5	AAU78269 Human tra
12	480	31.4	296	5	AD117131 Murine NO
13	474	31.0	1305	4	ABG27554 Novel hum
14	472	30.8	295	2	AAW23857 E6AP-bind
15	472	30.8	295	7	ADD18680 Human dis
16	472	30.8	295	8	ADP24302 PRO polyp
17	472	30.8	295	9	ADW71803 Human phe
18	467	30.5	296	5	AD117130 Rat NOVX
19	458	29.9	285	8	ADY80433 Rat aryl
20	458	28.9	291	7	ADA45448 Rat Prote
21	458	29.9	291	7	ADW57139 Rat Prote
22	458	29.9	291	7	ADW57135 Rat Prote
23	458	29.9	291	7	ADW57147 Rat Prote

97 357.5 23.4 254 8 ABM82695 Human dia  
98 357.5 23.4 254 8 ABM82697 Human dia  
99 357.5 23.4 283 7 ADD48500 Rat Prote  
100 356.5 23.3 285 10 AEF05774 Human hyd

ALIGNMENTS

RESULT 1

AAV79215  
ID AAV79215 standard; protein; 284 AA.

AC AAV79215;

DT 19-JUN-2000 (first entry)

DE Human transferase TRNSFS-7.

XX Transferase; TRNSFS-7; human; antitumour; cancer;  
KW gastrointestinal disorder; developmental disorder; genetic disorder;  
KW neurological disorder; reproductive disorder; smooth muscle disorder;  
KW immunological disorder; inflammation; diagnosis; therapy;  
KW sulfotransferase.

OS Homo sapiens.

XX Key Location/Qualifiers  
FH Modified-site 4  
FT Modified-site 11 /note= "potential O-phosphorylation"  
FT Modified-site 23..272 /note= "potential O-phosphorylation"  
FT Domain /note= "sulfotransferase protein domain"  
FT Modified-site 104 /note= "potential O-phosphorylation"  
FT Modified-site 148 /note= "potential O-phosphorylation"  
FT Modified-site 153 /note= "potential O-phosphorylation"  
FT Modified-site 157 /note= "potential O-phosphorylation"  
FT Modified-site 172 /note= "potential O-phosphorylation"  
FT Modified-site 174 /note= "potential O-phosphorylation"  
FT Modified-site 205 /note= "potential O-phosphorylation"  
FT Modified-site 260 /note= "potential O-phosphorylation"  
FT Modified-site /note= "potential O-phosphorylation"

XX WO200014251-A2.

XX 16-MAR-2000.

XX 09-SEP-1999; 99WO-US020989.

XX 10-SEP-1998; 98US-00150657.

XX 04-NOV-1998; 98US-00186779.

XX 11-MAY-1999; 99US-0133642P.

XX (INCY-) INCYTE PHARM INC.

XX Tang YT, Corley NC, Guegler KJ, Baughn MR, Lal P, Yue H;

XX Hillman JL, Azimzai Y;

XX WPI: 2000-256996/22.

XX N-PSDB; AA294207.

XX Human transferase proteins useful for preventing, diagnosing and treating  
PT cancers and developmental, gastrointestinal, genetic, immunological,  
PT neurological, reproductive and smooth muscle disorders.

PS Claim 1; Page 85-86; 113pp; English.

XX The present sequence is that of human transferase TRNSFS-7, 1 of 15  
CC claimed novel human transferase proteins of the invention (see AAY79209-  
CC 23). The sequence was deduced from a cDNA clone (see AA294207) isolated  
CC from a cerebellum library. It shows homology to sulfotransferase  
CC proteins. TRNSFS-7 is expressed in nervous, endocrine and development  
CC tissues, especially those associated with neurological disorders,  
CC inflammation and cancer. The new human transferases and polynucleotides  
CC can be used in the diagnosis, prevention and treatment of cancer,  
CC developmental disorders, gastrointestinal disorders, genetic disorders,  
CC immunological disorders, neurological disorders, reproductive disorders,  
CC and smooth muscle disorders. The polypeptides can also be used to raise  
CC antibodies, and to screen for agonists and antagonists of transferase  
CC activity

XX SQ Sequence 284 AA; #7

Query Match 100.0%; Score 1530; DB 3; Length 284;  
Best Local Similarity 100.0%; Pred. No. 2.3e-151;  
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAESEAEPTSPGFEFESKYFHFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60  
Db 1 MAESEAEPTSPGFEFESKYFHFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60  
QY 61 LQEVVYLVSGADPDEIGLMNIDBOLPVLEYVPOQGLDIIKELTSPRLIKSHLPVRFPSD 120  
Db 61 LQEVVYLVSGADPDEIGLMNIDBOLPVLEYVPOQGLDIIKELTSPRLIKSHLPVRFPSD 120  
QY 121 LHNGDSKVIYMARNPKDLVWSYQFHRSLRTWSYRGTFQECRRPMNDKLGSGWFHVQ 180  
Db 121 LHNGDSKVIYMARNPKDLVWSYQFHRSLRTWSYRGTFQECRRPMNDKLGSGWFHVQ 180  
QY 181 EFWHRMDSNVLFKYEDMHRDLVTWVQLARFLGVSCDKAQLAETECHQLVDPCCNA 240  
Db 181 EFWHRMDSNVLFKYEDMHRDLVTWVQLARFLGVSCDKAQLAETECHQLVDPCCNA 240  
QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284  
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 2

AAU78268

ID AAU78268 standard; protein; 284 AA.

AC AAU78268;

DT 05-JUN-2002 (first entry)

XX Human nervous system-specific sulfotransferase, SULTn, protein sequence.  
XX Human; nervous system-specific sulfotransferase; SULTn; sulfonation;  
KW neurological disorder; multiple sclerosis; Huntington's chorea;  
KW nervous system-mediated muscular atrophy; amyotrophic lateral sclerosis;  
KW myasthenia gravis; Alzheimer's disease; senile dementia; Schizophrenia;  
KW bipolar disorder; cytosolic sulfotransferase gene family;  
XX cerebellar ataxia; chromosome 22q13.

OS Homo sapiens.

XX Key Location/Qualifiers

FH Binding-site 54..61 /label= PAPS\_5'\_phosphosulfate\_binding\_site

FT Binding-site 129..144 /label= PAPS\_3'\_phosphate\_binding\_site

FT Binding-site 246..258 /label= PAPS\_3'\_phosphate\_binding\_site

XX WO200218541-A2.

XX 07-MAR-2002.

PD



XX 04-SEP-2001; 2001WO-US041986.  
 XX 01-SEP-2000; 2000US-0229929P.  
 XX (UYBO-) UNIV BOSTON.  
 XX Farb DH, Martin S;  
 XX WPI; 2002-281063/32.  
 XX N-PSDB; ABK12289.  
 XX New nervous system-specific sulfotransferase (SULTn) polypeptide for  
 PT treating neurological disorders including multiple sclerosis, nervous  
 PT system-mediated muscular atrophy and Huntington's chorea.  
 XX  
 PS Claim 2; Fig 1a; 46pp; English.  
 XX The present invention relates to a new nervous system-specific  
 CC sulfotransferase (SULTn) polypeptide. The SULTn polypeptide of the  
 CC invention comprises a nervous system-specific sulfotransferase,  
 CC comprising a sequence of 284 amino acids, given in the specification. The  
 CC invention is useful for identifying an effector of nervous system-  
 CC specific sulfonation by contacting the polypeptide with a candidate  
 CC effector compound and assaying the ability of the candidate effector  
 CC compound to effect the sulfotransferase function of the polypeptide. The  
 CC invention is also useful for inhibiting nervous system-specific  
 CC sulfonation by contacting the polypeptide under physiological conditions  
 CC with a compound capable of interacting with and inhibiting the  
 CC sulfonation activity of the polypeptide, where the compound is an  
 CC antibody, peptide, polypeptide, nucleic acid, organic molecule or an  
 CC inorganic molecule. The polypeptide of the invention is also useful for  
 CC treating a neurological disorder, including multiple sclerosis, nervous  
 CC system-mediated muscular atrophy, amyotrophic lateral sclerosis,  
 CC Huntington's chorea, myasthenia gravis, Alzheimer's disease, senile  
 CC dementia, schizophrenia, bipolar disorder and cerebellar ataxia. The  
 CC present amino acid sequence represents the human SULTn protein of the  
 CC invention. This sequence is encoded by the human SULTn gene located on  
 CC chromosome 22q13. SULTn is a member of the cytosolic sulfotransferase  
 CC gene family  
 XX  
 SQ Sequence 284 AA;  
 Query Match 100.0%; Score 1530; DB 5; Length 284;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-151;  
 Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60  
 Db 1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60  
 Qy 61 LQEVVYLVSQADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
 Db 61 LQEVVYLVSQADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
 Qy 121 LHNGDSKVIYNARNPKDLVSVYQPHRSRLTMSYRGTFQECRRPMNDKLGSGWFEHVQ 180  
 Db 121 LHNGDSKVIYNARNPKDLVSVYQPHRSRLTMSYRGTFQECRRPMNDKLGSGWFEHVQ 180  
 Qy 181 EFWEHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCKAQLEALTECHQLVDQCCNA 240  
 Db 181 EFWEHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCKAQLEALTECHQLVDQCCNA 240  
 Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284  
 Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284  
 RESULT 3  
 ID ABR01809  
 XX ABR01809 standard; protein; 284 AA.  
 AC ABR01809;

XX 19-MAY-2003 (first entry)  
 XX Human cancer-related protein, 192P2G7.  
 XX Human; cytostatic; vaccine; cancer; immune response.  
 XX Homo sapiens.  
 XX WO200283921-A2.  
 XX 24-OCT-2002.  
 XX 10-APR-2002; 2002WO-US011654.  
 XX 10-APR-2001; 2001US-0282739P.  
 XX 10-APR-2001; 2001US-0283112P.  
 XX 25-APR-2001; 2001US-0286630P.  
 XX (AGEN-) AGENSYS INC.  
 XX Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;  
 XX Morrison K, Morrison RK, Raitano AB;  
 XX WPI; 2003-075555/07.  
 XX N-PSDB; ABZ78140.  
 XX New composition comprising a substance that modulates the structure of  
 PT proteins and polynucleotides, useful for therapeutic, prognostic and  
 PT diagnostic reagents for eliciting cellular or humoral immune response in  
 PT cancer patients.  
 XX  
 PS Claim 12; Fig 2R; 1021pp; English.  
 XX The present invention relates to novel human cancer-related genes and  
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and  
 CC proteins are useful for eliciting a humoral or cellular immune response.  
 CC The genes are useful as probes and primers for the amplification and/or  
 CC detection of genes, mRNAs or their fragments, as reagents for the  
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
 CC directing the expression of the protein, as tools for modulating or  
 CC inhibiting the expression of genes and/or translation of transcripts, and  
 CC as therapeutic agents. The proteins and peptides are useful as  
 CC therapeutic, prognostic and diagnostic reagents for cancer  
 XX  
 SQ Sequence 284 AA;  
 Query Match 100.0%; Score 1530; DB 6; Length 284;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-151;  
 Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60  
 Db 1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60  
 Qy 61 LQEVVYLVSQADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
 Db 61 LQEVVYLVSQADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
 Qy 121 LHNGDSKVIYNARNPKDLVSVYQPHRSRLTMSYRGTFQECRRPMNDKLGSGWFEHVQ 180  
 Db 121 LHNGDSKVIYNARNPKDLVSVYQPHRSRLTMSYRGTFQECRRPMNDKLGSGWFEHVQ 180  
 Qy 181 EFWEHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCKAQLEALTECHQLVDQCCNA 240  
 Db 181 EFWEHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCKAQLEALTECHQLVDQCCNA 240  
 Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284  
 Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284  
 RESULT 4

```

ADQ89158
ID ADQ89158 standard; protein; 284 AA.
XX
AC ADQ89158;
XX
DT 21-OCT-2004 (first entry)
XX
DE Human urological disorder related protein 16386 SEQ:110.
XX
KW urological disorder; uropathic; cytostatic; urinary incontinence;
XX benign prostatic hyperplasia; human.
XX
OS Homo sapiens.
XX
PN WO2004065576-A2.
XX
PD 05-AUG-2004.
XX
PF 14-JAN-2004; 2004WO-US000750.
XX
PR 15-JAN-2003; 2003US-0440318P.
PR 04-FEB-2003; 2003US-0444783P.
PR 27-MAR-2003; 2003US-0457901P.
PR 08-MAY-2003; 2003US-0468775P.
PR 19-MAY-2003; 2003US-0471614P.
PR 16-JUN-2003; 2003US-0478742P.
PR 18-JUL-2003; 2003US-0488529P.
PR 30-JUL-2003; 2003US-0491156P.
PR 02-SEP-2003; 2003US-049594P.
PR 26-SEP-2003; 2003US-0506332P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Karicheti V, Silos-Santiago I, Eliasof SD;
XX
DR WPI; 2004-562167/54.
XX
DR N-PSDB; ADQ89157.
XX
PT Use of polypeptides related to urological disorders, e.g. 44390, 54181,
PT 211 or for identifying a compound capable of treating a urological
PT disorder or identifying and treating a subject having a urological
PT disorder.
XX
PS Claim 1; SEQ ID NO 110; 542pp; English.
XX
CC The present invention describes the use of polypeptides related to
CC urological disorders for identifying a compound capable of treating a
CC urological disorder, identifying a subject having a urological disorder,
CC or treating a subject having a urological disorder. Also described: (1) a
CC method for identifying a compound capable of treating a urological
CC disorder; (2) a method for identifying a subject having a urological
CC disorder; and (3) a method for treating a subject having a urological
CC disorder. The compound has uropathic and cytostatic activities. The
CC polypeptides related to urological disorders are useful for identifying a
CC compound capable of treating a urological disorder, identifying a subject
CC having a urological disorder, or treating a subject having a urological
CC disorder. Disorders include urinary incontinence and benign prostatic
CC hyperplasia. The present sequence represents a human urological disorder
CC related protein, which is used in the exemplification of the present
CC invention.
XX
SQ Sequence 284 AA;
Query Match 100.0%; Score 1530; DB 8; Length 284;
Best Local Similarity 100.0%; Pred. No. 2.3e-151;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAESEATPTSGFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVIWVTPKSGTSL 60
DB 1 MAESEATPTSGFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVIWVTPKSGTSL 60
QY 61 LQEVVYLSQCADPDEIGLNMNIDQLPVLEYPPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

```

```

Db 61 LQEVVYLSQCADPDEIGLNMNIDQLPVLEYPPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNPKOLVVSYYQFHRSLRTMSVGRGTFOEFCRRFMDKLGYSGWFEHVQ 180
DB 121 LHNGDSKVIYMARNPKOLVVSYYQFHRSLRTMSVGRGTFOEFCRRFMDKLGYSGWFEHVQ 180
QY 181 EFWEHRMDSNVLFLLKYEDMHRDLVTMVEQLARFLGVSCDKAQLALTEHCHQLVDQCCNA 240
DB 181 EFWEHRMDSNVLFLLKYEDMHRDLVTMVEQLARFLGVSCDKAQLALTEHCHQLVDQCCNA 240
QY 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
DB 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 5
ADR44884
ID ADR44884 standard; protein; 284 AA.
XX
AC ADR44884;
XX
DT 18-NOV-2004 (first entry)
XX
DE Polypeptide 16386 amino acid sequence.
XX
KW analgesic; gene therapy; pain; painful disorder.
XX
OS Homo sapiens.
XX
PN WO2004071411-A2.
XX
PD 26-AUG-2004.
XX
PF 30-JAN-2004; 2004WO-US002851.
XX
PR 04-FEB-2003; 2003US-0444781P.
PR 05-MAR-2003; 2003US-0452291P.
PR 13-MAR-2003; 2003US-0454540P.
PR 16-JUN-2003; 2003US-0478805P.
PR 30-JUL-2003; 2003US-0491048P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Silos-Santiago I, Karicheti V, Eliasof SD;
XX
DR WPI; 2004-625760/60.
XX
DR N-PSDB; ADR44883.
XX
PT Identifying a compound capable of treating pain by combining a compound
PT to be tested with e.g., 16386 polypeptide or with a host cell expressing
PT the polypeptide and detecting the binding of the test compound to the
PT polypeptide.
XX
PS Disclosure; SEQ ID NO 2; 233pp; English.
XX
CC The invention relates to a method of identifying a compound capable of
CC treating pain or painful disorder by combining a compound to be tested
CC with an e.g., 16386, 15402, 21165, 1423, 636, or 32620 polypeptide or with
CC a host cell expressing the polypeptide and detecting the binding of the
CC test compound to the polypeptide to identify a compound that binds to the
CC polypeptide. The method is useful in identifying a compound capable of
CC treating pain or painful disorder. This sequence corresponds to the
CC polypeptide 16386 - a brain sulfoltransferase-like protein.
XX
SQ Sequence 284 AA;
Query Match 100.0%; Score 1530; DB 8; Length 284;
Best Local Similarity 100.0%; Pred. No. 2.3e-151;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAESEATPTSGFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVIWVTPKSGTSL 60
DB 1 MAESEATPTSGFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVIWVTPKSGTSL 60

```

Qy 61 LOEVVYLSQGADPDEIGLMNIDQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Db 61 LOEVVYLSQGADPDEIGLMNIDQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Qy 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRPMNDKLGYSWFHEVQ 180  
Db 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRPMNDKLGYSWFHEVQ 180  
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240  
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240  
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284  
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284

RESULT 6  
ADZ211933  
ID ADZ211933 standard; protein; 284 AA.  
XX  
AC ADZ211933;  
XX  
DT 16-JUN-2005 (first entry)  
XX  
DE Human SUL4A1 splice variant 1 SEQ ID NO 1.  
XX  
KW neuroprotective; nootropic; diagnosis; prognosis; gene expression;  
KW neurodegenerative disease; neurological disease; Alzheimers disease;  
KW degeneration; sulfotransferase family 4A member 1; SUL4A1;  
KW splice variant; enzyme.  
XX  
OS Homo sapiens.  
XX  
PN WO2005030947-A2.  
XX  
PD 07-APR-2005.  
XX  
PF 29-SEP-2004; 2004WO-EP052353.  
XX  
PR 30-SEP-2003; 2003US-0506775P.  
XX  
PA (EVOT-) EVOTEC NEUROSCIENCES GMBH.  
XX  
PI Von Der Kammer H, Pohlner J;  
XX  
XX WPI; 2005-262951/27.  
DR N-PSDB; ADZ21938, ADZ21935.  
XX  
XX

Diagnosing or prognosticating a neurodegenerative disease in a subject by determining a level and/or activity of a transcription and/or translation product of a gene coding for a cytosolic sulfotransferase family 4A

Claim 15; SEQ ID NO 1; 83pp; English.

The invention describes a method of diagnosing or prognosticating a neurodegenerative disease in a subject, or determining whether a subject is at increased risk of developing the disease. The method comprises determining a level and/or an activity of a transcription product and/or a translation product of a gene coding for a cytosolic sulfotransferase family 4A member 1 and/or a fragment, or derivative or variant of the transcription or translation product, in a sample obtained from the subject and comparing the level and/or the activity to a reference value representing a known disease or health status, therefore diagnosing or prognosticating the neurodegenerative disease in the subject, or determining whether the subject is at increased risk of developing the neurodegenerative disease. Protein molecules comprising fully defined 284 and/or 171 amino acid (SEQ ID NO. 1 and/or 2, respectively) sequence given in the specification, the protein molecules being translation products of the gene coding for a cytosolic sulfotransferase family 4A member 1, or its fragments, or derivatives, or variants, are useful as

CC diagnostic targets for detecting a neurodegenerative disease, preferably Alzheimer's disease or as screening targets for reagents or compounds preventing, or treating, or ameliorating the disease. An antibody specifically immunoreactive with an immunogen, where the immunogen is a translation product of a gene coding for a cytosolic sulfotransferase family 4A member 1, SEQ ID NO. 1 or 2, or its fragment, or derivative, or variant is useful for detecting the pathological state of a cell in a sample obtained from a subject, comprising immunocytochemical staining of the cell with the antibody, where an altered degree of staining, or an altered staining pattern in the cell compared to a cell representing a known health status indicates a pathological state of the cell which relates to a neurodegenerative disease, preferably to Alzheimer's disease. The genetically altered non-human animal is useful for screening, testing, and validating compounds, agents, and modulators in the development of diagnostics and therapeutics to treat neurodegenerative diseases, in particular Alzheimer's disease. An agent(s) which directly or indirectly affect an activity and/or a level of a gene coding for a cytosolic sulfotransferase family 4A member 1, and/or a transcription product or a translation product of a gene coding for a cytosolic sulfotransferase family 4A member 1, and/or their fragment, or derivative, or variant is useful for treating or preventing a neurodegenerative disease, in particular Alzheimer's disease. This is the amino acid sequence of human sulfotransferase family 4A member 1 splice variant 1 (SULT4A1sv1).

XX  
SQ Sequence 284 AA;

Query Match 100.0%; Score 1530; DB 9; Length 284;  
Best Local Similarity 100.0%; Pred. No. 2.3e-151;  
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEATPTSGEFESKYFEFHGVRLPPFCRGKMEBIANPPVPRPSDVWIVTPKSGTSL 60  
Db 1 MAESEATPTSGEFESKYFEFHGVRLPPFCRGKMEBIANPPVPRPSDVWIVTPKSGTSL 60  
Qy 61 LOEVVYLSQGADPDEIGLMNIDQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Db 61 LOEVVYLSQGADPDEIGLMNIDQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Qy 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRPMNDKLGYSWFHEVQ 180  
Db 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRPMNDKLGYSWFHEVQ 180  
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240  
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240  
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284  
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284

RESULT 7  
ABM04807  
ID ABM04807 standard; protein; 284 AA.  
XX  
XX AC ABM04807;  
XX  
XX DT 22-SEP-2003 (first entry)  
XX  
XX DE Rat sulfotransferase-like protein.  
XX  
XX KW spinal cord; neuropathic pain; central sensitisation pain; pain;  
XX KW analgesic; gene therapy.  
XX  
XX OS Rattus norvegicus.  
XX  
XX PN EP1284298-A2.  
XX  
XX PD 19-FEB-2003.  
XX  
XX PF 26-JUL-2002; 2002EP-00255229.  
XX

```

PR 27-JUL-2001; 2001GB-00018354.
PR 07-FEB-2002; 2002GB-00002883.
XX
XX (WARN ) WARNER LAMBERT CO.
XX Brooksbank RA, Dixon AK, Lee K, Pincock RD;
XX WPI; 2003-543489/52.
XX N-P8DB; ACF25349.
XX
XX Use of an isolated gene sequence in the screening of compounds for
XX diagnosing or treating pain.
XX Claim 1; Page 93-94; 188pp; English.
XX
XX The invention relates to a novel isolated gene sequence that is
XX downregulated in the spinal cord of a mammal in response to mechanically
XX distinct first and second models of neuropathic or central sensitisation
XX pain, useful in the screening of compounds for diagnosing or treating
XX pain. A protein encoded by a gene of the invention has analgesic
XX activity. A polynucleotide of the invention may have a use in gene
XX therapy. The gene sequence is useful for preparing a composition for
XX diagnosing or treating pain. The present sequence represents a protein
XX encoded by a gene of the invention
XX
SQ Sequence 284 AA;
Query Match 98.4%; Score 1506; DB 6; Length 284;
Best Local Similarity 97.9%; Pred. No. 7.4e-149;
Matches 278; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MAESEAEPTSPGPFESKYFPHGVRLPPFCRGKMEETANFPVPSDVWIVTPKSGTSL 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MAESEAEPTGPGFESKYFPHGVRLPPFCRGKMEETADFPVPSDVWIVTPKSGTSL 60
Qy 61 LQEVVYLVSGADPDEIGLMNIDEOLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 LQEVVYLVSGADPDEIGLMNIDEOLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYMARNPKDLVSYQFHRSLRTMSYRGTFQECRRFMDKLGYSWFEHVQ 180
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 LHNGDSKVIYMARNPKDLVSYQFHRSLRTMSYRGTFQECRRFMDKLGYSWFEHVQ 180
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLTEHCHQLVDQCNA 240
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLTEHCHQLVDQCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSNNEKFDLVYKQKMGKCDLTDFDYL 284
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 EALPVGRGRVGLWKDIFTVSNNEKFDLVYKQKMGKCDLTDFDYL 284

RESULT 8
ABR01860
ID ABR01860 standard; protein; 218 AA.
XX
XX ABR01860;
XX
XX 19-MAY-2003 (first entry)
XX Human cancer-related protein, 192P2G7 v.2.
XX Human; cytostatic; vaccine; cancer; immune response.
XX Homo sapiens.
XX WO200283921-A2.
XX 24-OCT-2002.
XX
XX 10-APR-2002; 2002WO-US011654.
XX
XX 10-APR-2001; 2001US-0282739P.

```

```

PR 10-APR-2001; 2001US-0283112P.
PR 25-APR-2001; 2001US-0286630P.
XX
XX (AGEN-) AGENSYS INC.
XX Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
XX Morrison K, Morrison RK, Raitano AB;
XX
XX WPI; 2003-075555/07.
XX N-PSDB; ABZ78167.
XX
XX New composition comprising a substance that modulates the structure of
XX proteins and polynucleotides, useful for therapeutic, prognostic and
XX diagnostic reagents for eliciting cellular or humoral immune response in
XX cancer patients.
XX Example 5; Fig 12r; 1021pp; English.
XX
XX The present invention relates to novel human cancer-related genes and
XX proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
XX proteins are useful for eliciting a humoral or cellular immune response.
XX The genes are useful as probes and primers for the amplification and/or
XX detection of genes, mRNAs or their fragments, as reagents for the
XX diagnosis and/or prognosis of cancer, as coding sequences capable of
XX directing the expression of the protein, as tools for modulating or
XX inhibiting the expression of genes and/or translation of transcripts, and
XX as therapeutic agents. The proteins and peptides are useful as
XX therapeutic, prognostic and diagnostic reagents for cancer
XX
SQ Sequence 218 AA;
Query Match 65.5%; Score 1002; DB 6; Length 218;
Best Local Similarity 56.4%; Pred. No. 3.6e-96;
Matches 186; Conservative 0; Mismatches 1; Indels 6; Gaps 1;
Qy 92 PQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVSYQFHRSLRT 151
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
32 PSP-----ELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVSYQFHRSLRT 85
Qy 152 MSYRGTFQECRRFMDKLGYSWFEHVQFWEHRMDSNVLFKYEDMHRDLVTWVEQLA 211
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
86 MSYRGTFQECRRFMDKLGYSWFEHVQFWEHRMDSNVLFKYEDMHRDLVTWVEQLA 145
Qy 212 RFLGVSCDKAQLTEHCHQLVDQCNAEALPVGRGRVGLWKDIFTVSNNEKFDLVYKQ 271
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
146 RFLGVSCDKAQLTEHCHQLVDQCNAEALPVGRGRVGLWKDIFTVSNNEKFDLVYKQ 205
Qy 272 KMGKCDLTDFDYL 284
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
206 KMGKCDLTDFDYL 218

RESULT 9
ABR01861
ID ABR01861 standard; protein; 171 AA.
XX
XX ABR01861;
XX
XX 19-MAY-2003 (first entry)
XX Human cancer-related protein, 192P2G7 v.3.
XX Human; cytostatic; vaccine; cancer; immune response.
XX Homo sapiens.
XX WO200283921-A2.
XX 24-OCT-2002.
XX
XX 10-APR-2002; 2002WO-US011654.
XX
XX 10-APR-2001; 2001US-0282739P.

```

PR	10-APR-2001; 2001US-0283112P.	PN	WO2005030947-A2.
PR	25-APR-2001; 2001US-0286630P.	XX	
XX		PD	07-APR-2005.
PA	(AGEN-) AGENSYS INC.	XX	
XX		PF	29-SEP-2004; 2004WO-EP052353.
XX	Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;	XX	
PI	Morrison K, Morrison RK, Raitano AB;	PR	30-SEP-2003; 2003US-0506775P.
XX		XX	(EVOT-) EVOTEC NEUROSCIENCES GMBH.
DR	WPI; 2003-075555/07.	PA	
DR	N-PSDB; ABZ78168.	XX	
XX		PI	Von Der Kammer H, Pohlner J;
XX	New composition comprising a substance that modulates the structure of	XX	
PT	proteins and polynucleotides, useful for therapeutic, prognostic and	DR	WPI; 2005-262951/27.
PT	diagnostic reagents for eliciting cellular or humoral immune response in	DR	N-PSDB; ADZ21936.
PT	cancer patients.	XX	
XX		PT	Diagnosing or prognosticating a neurodegenerative disease in a subject by
XX	Example 5; Fig 12r; 1021pp; English.	PT	determining a level and/or activity of a transcription and/or translation
XX		PT	product of a gene coding for a cytosolic sulfotransferase family 4A
PS		PT	member 1.
CC	The present invention relates to novel human cancer-related genes and	CC	Claim 15; SEQ ID NO 2; 83pp; English.
CC	proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and	XX	
CC	proteins are useful for eliciting a humoral or cellular immune response.	PS	
CC	The genes are useful as probes and primers for the amplification and/or	XX	
CC	detection of genes, mRNAs or their fragments, as reagents for the	CC	The invention describes a method of diagnosing or prognosticating a
CC	diagnosis and/or prognosis of cancer, as coding sequences capable of	CC	neurodegenerative disease in a subject, or determining whether a subject
CC	directing the expression of the protein, as tools for modulating or	CC	is at increased risk of developing the disease. The method comprises
CC	inhibiting the expression of genes and/or translation of transcripts, and	CC	determining a level and/or an activity of a transcription product and/or
CC	as therapeutic agents. The proteins and peptides are useful as	CC	a translation product of a gene coding for a cytosolic sulfotransferase
CC	therapeutic, prognostic and diagnostic reagents for cancer	CC	family 4A member 1 and/or a fragment, or derivative or variant of the
XX		CC	transcription or translation product, in a sample obtained from the
XX	Sequence 171 AA;	CC	subject and comparing the level and/or the activity to a reference value
QQ		CC	representing a known disease or health status, therefore diagnosing or
	Query Match 56.7%; Score 867.5; DB 6; Length 171;	CC	prognosticating the neurodegenerative disease in the subject, or
	Best Local Similarity 59.9%; Pred. No. 3.2e-82;	CC	determining whether the subject is at increased risk of developing the
	Matches 170; Conservative 1; Mismatches 0; Indels 113; Gaps 1;	CC	neurodegenerative disease. Protein molecules comprising fully defined 284
QY	1 MAESEAETSTGPEPSKFEFHGVRLLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 60	CC	and/or 171 amino acid (SEQ ID NO. 1 and/or 2, respectively) sequence
Db		CC	given in the specification, the protein molecules being translation
	1 MAESEAETSTGPEPSKFEFHGVRLLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 56	CC	products of the gene coding for a cytosolic sulfotransferase family 4A
		CC	member 1, or its fragments, or derivatives, or variants, are useful as
QY	61 LQEVVVLVSGQADPDIGLWNIDBQLPVLEYQPQGLDIIKELTSPRLIKSLHPYRPLPSD 120	CC	diagnostic targets for detecting a neurodegenerative disease, preferably
Db		CC	Alzheimer's disease or as screening targets for reagents or compounds
	57 -----	CC	preventing, or treating, or ameliorating the disease. An antibody
	57 -----	CC	specifically immunoreactive with an immunogen, where the immunogen is a
QY	121 LHNGDSKVITYMARNPKDLVVSYQFHRSLRTWSYRGTFQFCRRFMNDKLGYSWFHVQ 180	CC	translation product of a gene coding for a cytosolic sulfotransferase
Db		CC	family 4A member 1, SEQ ID NO. 1 or 2, or its fragment, or derivative, or
	57 -----	CC	variant is useful for detecting the pathological state of a cell in a
QY	181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARPLGVSCDKAQLTEALTECHQLVDQCCNA 240	CC	sample obtained from a subject, comprising immunocytochemical staining of
Db		CC	the cell with the antibody, where an altered degree of staining or an
	68 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARPLGVSCDKAQLTEALTECHQLVDQCCNA 127	CC	altered staining pattern in the cell compared to a cell representing a
		CC	known health status indicates a pathological state of the cell which
QY	241 EALPVGRGVLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFVL 284	CC	relates to a neurodegenerative disease, preferably to Alzheimer's
Db		CC	disease. The genetically altered non-human animal is useful for
	128 EALPVGRGVLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFVL 171	CC	screening, testing, and validating compounds, agents, and modulators in
		CC	the development of diagnostics and therapeutics to treat
RESULT 10		CC	neurodegenerative diseases, in particular Alzheimer's disease. An
ID	ADZ21934	CC	agent(s) which directly or indirectly affect an activity and/or a level
XX	ADZ21934 standard; protein; 171 AA.	CC	of a gene coding for a cytosolic sulfotransferase family 4A member 1,
AC	ADZ21934;	CC	and/or a transcription product or a translation product of a gene coding
DT	16-JUN-2005 (first entry)	CC	for a cytosolic sulfotransferase family 4A member 1, and/or their
XX		CC	fragment, or derivative, or variant is useful for treating or preventing
DE	Human SULT4A1 splice variant 2 SEQ ID NO 2.	CC	a neurodegenerative disease, in particular Alzheimer's disease. This is
XX		CC	the amino acid sequence of human sulfotransferase family 4A member 1
KW	neuroprotective; nootropic; diagnosis; prognosis; gene expression;	CC	splice variant 2 (SULT4A1svl).
KW	neurodegenerative disease; neurological disease; Alzhemiers disease;	XX	Sequence 171 AA;
KW	degeneration; sulfotransferase family 4A member 1; SULT4A1;	Query Match 56.7%; Score 867.5; DB 9; Length 171;	
XX	splice variant; enzyme.	Best Local Similarity 59.9%; Pred. No. 3.2e-82;	
OS	Homo sapiens.	Matches 170; Conservative 1; Mismatches 0; Indels 113; Gaps 1;	
XX		QY	1 MAESEAETSTGPEPSKFEFHGVRLLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 60
		Db	
			1 MAESEAETSTGPEPSKFEFHGVRLLPPFCRGMKEEIANFPVRPSDVWIVTPKSGTSL 56

QY 61 LQEVVYLVSQGADPDEIGLWNIDEQLPVLEYPQPGLDIIKELTSPRLIKSHLPYRFLPSD 120  
 Db 57 ----- 56  
 QY 121 LHNGSKVIYMARNPDLVSYQPHRSRLTMSYRGTFQEFRCRPMNDKLGYSWFHVQ 180  
 Db 57 -----VGYSWFHVQ 67  
 QY 181 EFWHRMDSNVLFKLYEDMHRDLVTMVQLARFLGVSCDKAQLALETCHQHVLDQCCNA 240  
 Db 68 EFWHRMDSNVLFKLYEDMHRDLVTMVQLARFLGVSCDKAQLALETCHQHVLDQCCNA 127  
 QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284  
 Db 128 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 171

RESULT 11  
 AAU78269  
 ID AAU78269 standard; protein; 159 AA.  
 XX  
 AC AAU78269;  
 DT 05-JUN-2002 (first entry)  
 XX  
 DE Rat nervous system-specific sulfotransferase, SULTn, protein sequence.  
 XX  
 KW Rat; nervous system-specific sulfotransferase; SULTn; sulfonation;  
 KW neurological disorder; multiple sclerosis; Huntington's chorea;  
 KW nervous system-mediated muscular atrophy; amyotrophic lateral sclerosis;  
 KW myasthenia gravis; Alzheimer's disease; senile dementia; Schizophrenia;  
 KW bipolar disorder; cytosolic sulfotransferase gene family;  
 KW cerebellar ataxia.  
 XX  
 OS Rattus sp.  
 XX  
 PN WO200218541-A2.  
 XX  
 PD 07-MAR-2002.  
 XX  
 PF 04-SEP-2001; 2001WO-US041986.  
 XX  
 PR 01-SEP-2000; 2000US-0229929P.  
 XX  
 PA (UYBO-) UNIV BOSTON.  
 XX  
 PI Farb DH, Martin S;  
 XX  
 DR WPI; 2002-281063/32.  
 XX  
 DR P-PSDB; AAU78269.  
 XX  
 PT New nervous system-specific sulfotransferase (SULTn) polypeptide for  
 PT treating neurological disorders including multiple sclerosis, nervous  
 PT system-mediated muscular atrophy and Huntington's chorea.  
 XX  
 PS Claim 2; Page; 46pp; English.  
 XX

The present invention relates to a new nervous system-specific  
 CC sulfotransferase (SULTn) polypeptide. The SULTn polypeptide of the  
 CC invention comprises a nervous system-specific sulfotransferase,  
 CC comprising a sequence of 284 amino acids, given in the specification. The  
 CC invention is useful for identifying an effector of nervous system-  
 CC specific sulfonation by contacting the polypeptide with a candidate  
 CC effector compound and assaying the ability of the candidate effector  
 CC compound to affect the sulfotransferase function of the polypeptide. The  
 CC invention is also useful for inhibiting nervous system-specific  
 CC sulfonation by contacting the polypeptide under physiological conditions  
 CC with a compound capable of interacting with and inhibiting the  
 CC sulfonation activity of the polypeptide, where the compound is an  
 CC antibody, peptide, polypeptide, nucleic acid, organic molecule or an  
 CC inorganic molecule. The polypeptide of the invention is also useful for  
 CC treating a neurological disorder, including multiple sclerosis, nervous  
 CC system-mediated muscular atrophy, amyotrophic lateral sclerosis,

CC Huntington's chorea, myasthenia gravis, Alzheimer's disease, senile  
 CC dementia, Schizophrenia, bipolar disorder and cerebellar ataxia. The  
 CC present amino acid sequence represents the rat SULTn protein of the  
 CC family. SULTn is a member of the cytosolic sulfotransferase gene  
 CC family. Note: The sequence data for this sequence did not form part of  
 CC the printed specification but was obtained in electronic format from  
 CC GENBANK AF176343  
 XX  
 SQ Sequence 159 AA;  
 Query Match 55.6%; Score 850.5; DB 5; Length 159;  
 Best Local Similarity 98.1%; Pred. No. 1.7e-80;  
 Matches 157; Conservative 2; Mismatches 0; Indels 1; Gaps 1;  
 QY 125 DSKVIYMARNPDLVSYQPHRSRLTMSYRGTFQEFRCRPMNDKLGYSWFHVQEFWE 184  
 Db 1 DSKVIYMARNPDLVSYQPHRSRLTMSYRGTFQEFRCRPMNDKLGYSWFHVQEFWE 60  
 QY 185 HRMDSNVLFKLYEDMHRDLVTMVQLARFLGVSCDKAQLALETCHQHVLDQCCNAEALP 244  
 Db 61 HRMDANVLFKLYEDMHRDLVTMVQLARFLGVSCDKAQLALETCHQHVLDQCCNAEALP 119  
 QY 245 VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284  
 Db 120 VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 159

RESULT 12  
 ADI17131  
 ID ADI17131 standard; protein; 296 AA.  
 XX  
 AC ADI17131;  
 XX  
 DT 15-APR-2004 (first entry)  
 XX  
 DE Murine NOVX protein homologue SeqID 667.  
 XX  
 KW mouse; murine; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;  
 KW inflammation; autoimmune disorder; allergy; blood disorder;  
 KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;  
 KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;  
 KW Alzheimer's disease; infect.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO200268649-A2.  
 XX  
 PD 06-SEP-2002.  
 XX  
 PF 31-JAN-2002; 2002WO-US002785.  
 XX  
 PR 31-JAN-2001; 2001US-0265395P.  
 PR 31-JAN-2001; 2001US-0265412P.  
 PR 31-JAN-2001; 2001US-0265514P.  
 PR 02-FEB-2001; 2001US-0265517P.  
 PR 02-FEB-2001; 2001US-0266406P.  
 PR 05-FEB-2001; 2001US-0266767P.  
 PR 07-FEB-2001; 2001US-0266975P.  
 PR 07-FEB-2001; 2001US-0267057P.  
 PR 08-FEB-2001; 2001US-0267459P.  
 PR 09-FEB-2001; 2001US-0267823P.  
 PR 15-FEB-2001; 2001US-0268974P.  
 PR 26-FEB-2001; 2001US-0271664P.  
 PR 27-FEB-2001; 2001US-0271839P.  
 PR 27-FEB-2001; 2001US-0271855P.  
 PR 02-MAR-2001; 2001US-0272788P.  
 PR 02-MAR-2001; 2001US-0273046P.  
 PR 14-MAR-2001; 2001US-0275925P.  
 PR 14-MAR-2001; 2001US-0275947P.  
 PR 14-MAR-2001; 2001US-0275950P.  
 PR 14-MAR-2001; 2001US-0275989P.  
 PR 15-MAR-2001; 2001US-0276448P.  
 PR 15-MAR-2001; 2001US-0276450P.



CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostic, forensic, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 1305 AA;

Query Match 31.0%; Score 474; DB 4; Length 1305;  
Best Local Similarity 33.5%; Pred. No. 1e-39;  
Matches 110; Conservative 54; Mismatches 110; Indels 54; Gaps 8;

QY 3 BSEATPTSP-----GEFESKYFEFHGVL-----PP--FCRG-----K 34  
DB 265 EGRQPYTPYTPILSPCRQENSGSRELNMELIOTSRPPLYKGVPLIKYFAEA 324  
QY 35 MEEIANFVRPSDVIVTPKSGTSLQEVVYLVSGADPDEIGLNMIDQLPVLEYPQP 94  
DB 325 LGPLOSFOARPDDLINTYPKSGTTWVSQILDIMYQGGDLEKCNRAPIYVRVPFLEVNDP 384  
QY 95 ---GLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVITYMARNPKDLVVSYYQFHSRL 150  
DB 385 GPESGLETLDKTPPRLIKSHLPALLPQTLDDQKVVVYVARNPKDVAVSYHPRMEK 444  
QY 151 TMSYRGTFQFCRRFMNDKLGYSWFEHVEHVRMDSNVLFKLYEDMHRDLVTMVQOL 210  
DB 445 AHPEGTWDSFLEKFMAGEVSGSYQHVQEWELSRTHPVLYLYFEDMKNPKREIQKI 504  
QY 211 ARFLGVSCDKAQLALTEHC-----HQLVQCCNAEALP-VGRGRVGL 252  
DB 505 LEFVGRSLPEETDMFMVQHTSFKETKKNPMTNTYTPQELMDHSIS----PFMRKGMAGD 560  
QY 253 WKDIETVSMNEKFDLYVKQKGCIDLTF 280  
DB 561 WKTTFVAQNERFDADYAEKMGAGCSLSF 588

RESULT 14  
AAW23657  
ID AAW23657 standard; protein; 295 AA.  
XX  
AC AAW23657;  
XX  
DT 10-OCT-1997 (first entry)  
XX  
DE E6AP-binding protein cln25.  
XX  
KW Human; cell differentiation; survival; carcinoma; psoriasis;  
KW ubiquitination; p53; tumour suppressor; homeostasis; papilloma virus;  
KW epithelial cell; acne; ichthyosis; aphthous ulcer; hair growth; antibody;  
KW cell proliferation.  
XX  
OS Homo sapiens.  
XX  
PN WO9640767-A2.  
XX  
PD 19-DEC-1996.  
XX  
XX 06-JUN-1996; 96WO-US009040.  
XX  
PR 07-JUN-1995; 95US-00484878.  
XX

PA (MITO-) MITOTIX INC.  
XX Bear-Romero PL, Draetta G, Rolfe M;  
XX WPI; 1997-087053/08.  
DR N-PSDB; AAT78309.  
XX  
PT E6AP-binding proteins and related nucleic acid - useful for modulating  
PT cell differentiation, survival etc., partic. for treatment and diagnosis  
PT of carcinoma, psoriasis, etc.  
XX  
PS Claim 1; Page 70-71; 83pp; English.  
XX  
CC The present sequence represents the E6AP-binding protein cln25. E6AP  
CC mediates ubiquitination and so the inactivation of e.g. p53 tumour  
CC suppressor. The E6AP-binding protein is likely to be involved in normal  
CC cell homeostasis and in the pathogenesis of proliferative and  
CC differentiation disorders, e.g. regulation of gene expression or the cell  
CC cycle, modification of cell surface receptors, biogenesis of ribosomes  
CC and DNA repair. The protein, which can optionally be generated in vivo by  
CC gene therapy, may be useful in treatment and prevention of papilloma virus  
CC infected transformed cells and carcinoma, and may also be used to  
CC regulate epithelial cell processes more generally, e.g. in cases of  
CC psoriasis, acne, ichthyosis, aphthous ulcers. It may also be used to  
CC control wound healing, and inhibit growth of hair. It can also be used to  
CC generate antibodies which are used in immunoassays to determine the  
CC protein levels. The encoding nucleic acid can be used to prepare  
CC recombinant proteins and oligonucleotides useful as probes and primers  
CC for diagnostic detection of (mutant) mRNA for the protein in  
CC (transformed) cells, also for antisense therapy. It can also be used for  
CC the detection of mutations in E6AP-binding protein-encoding genes, mis-  
CC expression of these genes or mis-incorporation of the protein in a  
CC transcription regulatory complex containing E6AP can be used to assess  
CC risk of disorders characterised by cell proliferation  
XX  
SQ Sequence 295 AA;

Query Match 30.8%; Score 472; DB 2; Length 295;  
Best Local Similarity 35.8%; Pred. No. 2e-40;  
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLPFPFCRGKMERIANFVRPSDVIVTPKSGTSLQEVVYLVSGADPDEIGLNMID 83  
DB 17 GVPLIKYFAEALGPLQSFQARPDDLINTYPKSGTTWVSQILDIMYQGGDLEKCNRAPIY 76  
QY 84 EQLPVLEYPQP-----GLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVITYMARNPKDLV 139  
DB 77 VRVPFLEVNDPEFSGLETLDKTPPRLIKSHLPALLPQTLDDQKVVVYVARNPKDVA 136  
QY 140 VSYQFHRSLRMTSYRGTFQFCRRFMNDKLGYSWFEHVEHVRMDSNVLFKLYEDM 199  
DB 137 VSYHFRMEKAHPDPGTWDSFLEKFMAGEVSGSYQHVQEWELSRTHPVLYLYFEDM 196  
QY 200 HRDLVTMVQELARFLGVSCDKAQLALTEHC-----HQLVQCCNAE 242  
DB 197 KENPKREIQKILEFVGRSLPEETDMFMVQHTSFKEMKKNPMTNTYTPQELMDHSIS--- 253  
QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLYVKQKGCIDLTF 280  
DB 254 -PFMRKGMAGDKTTFVAQNERFDADYAEKMGAGCSLSF 291

RESULT 15  
ADD18680  
ID ADD18680 standard; protein; 295 AA.  
XX  
AC ADD18680;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Human disease related protein SeqID111.  
XX  
KW human; disease state; cytostatic; antiinflammatory; ophthalmological;



KW antiarteriosclerotic; vulnary; gene therapy;  
KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;  
KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;  
KW glucose transportation; catecholamine synthesis; iron transport;  
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;  
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;  
KW inflammatory condition; wound healing.  
XX  
OS Homo sapiens.  
XX  
XX WO2003018621-A2.  
XX  
XX PD 06-MAR-2003.  
XX  
XX PF 23-AUG-2002; 2002WO-GB003892.  
XX  
XX PR 23-AUG-2001; 2001GB-00020558.  
XX  
XX PR 05-OCT-2001; 2001GB-00024037.  
XX  
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.  
XX  
XX PI Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;  
XX  
XX DR WPI: 2003-290046/28.  
XX  
XX DR N-PSDB; ADD18681.  
XX  
XX PT New substantially purified polypeptide, useful for diagnosing or treating  
XX PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion  
XX PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or  
XX PT wound healing.  
XX  
XX PS Claim 25; SEQ ID NO 111; 424pp; English.  
XX  
XX CC This invention relates to novel human genes and gene product which are  
XX CC implicated in certain disease states. Compounds which modulate the  
XX CC proteins of the invention may have cytostatic, antiinflammatory,  
XX CC ophthalmological, antiarteriosclerotic or vulnary activities. The  
XX CC sequences of the invention may be useful for gene therapy. The invention  
XX CC may be useful for diagnosing or treating a hypoxia-regulated condition,  
XX CC such as tumorigenesis, angiogenesis, apoptosis, inflammation,  
XX CC erythropoiesis, or the biological response to hypoxia conditions  
XX CC including processes such as glycolysis, gluconeogenesis, glucose  
XX CC transportation, catecholamine synthesis, iron transport or nitric oxide  
XX CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion  
XX CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,  
XX CC inflammatory conditions or wound healing. The present sequence is that of  
XX CC a disease related protein of the invention.  
XX  
XX SQ Sequence 295 AA;  
Query Match 30.8%; Score 472; DB 7; Length 295;  
Best Local Similarity 35.8%; Pred. No. 2e-40;  
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;  
QY 24 GVRLPFFCRGKMEETANFVRPSDVWIVTPKSGTSLQEVYLVSQGADPEIGLMNID 83  
DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
17 GVPLIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDMYQGDLEKCNRAPIY 76  
QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
84 EQLPVLEYQPQ----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYVARNPKDLV 139  
DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
77 VRVPFLEVNDGPEPSGLTLDTPPPRLIKSHLPALLPQTLQDQKVVYVARNPKDVA 136  
QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
140 VSYQFHRSLRTMSYRGTFQBFRCRRFMNDKLGYSWFHVFQFWEHRMDSNVFLFKYEDM 199  
DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
137 VSYVHFHMEKXAHPEPGTWSDFLEKFMAGEVSYGSWYQVHVEWELSTHFPVLYLFYEDM 196  
QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
200 HRDLVTWVEQLARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAA 242  
DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
197 KENPKREIQKILEFVGRSLPSETMDFVQHTSFKEMKKNPMNTYTPVQELMDHSIS--- 253  
QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
243 LP-VGRGVGLWKLDFVTSMNEKEDLVYKQMGKCDLTF 280  
DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
254 -PFMRKGNAGDWKTTFTTVAQNERFDADYAEKMGAGCSLSF 291

## RESULT 16

ADP24302  
ID ADP24302 standard; protein; 295 AA.  
XX  
XX AC ADP24302;  
XX  
XX DT 18-NOV-2004 (first entry)  
XX  
XX DE PRO polypeptide SEQ ID NO:1480.  
XX  
XX KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;  
KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;  
KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.  
XX  
XX OS Unidentified.  
XX  
XX PN WO2004041170-A2.  
XX  
XX PD 21-MAY-2004.  
XX  
XX PF 30-OCT-2003; 2003WO-US034312.  
XX  
XX PR 01-NOV-2002; 2002US-0423394P.  
XX  
XX PA (GETH ) GENENTECH INC.  
XX  
XX PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;  
PI Wu TD;  
XX  
XX DR WPI; 2004-419628/39.  
XX  
XX DR N-PSDB; ADP24301.  
XX  
XX PT New PRO polypeptides and polynucleotides, useful for treating e.g.  
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated  
PT renal disease, or demyelinating diseases of the central or peripheral  
XX nervous system.  
XX  
XX PS Claim 7; SEQ ID NO 1480; 2940pp; English.  
XX  
XX CC The invention relates to a novel isolated nucleic acid and the PRO  
XX CC polypeptide encoded by it. A protein of the invention has  
XX CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,  
XX CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,  
XX CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide  
XX CC of the invention may have a use in gene therapy. The PRO polypeptide, its  
XX CC agonist, antagonist, or antibody that specifically binds to the  
XX CC polypeptide is useful for treating an immune related disorder such as  
XX CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,  
XX CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
XX CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic  
XX CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
XX CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
XX CC disease, a demyelinating disease of the central or peripheral nervous  
XX CC system, idiopathic demyelinating polynuropathy, Guillain-Barre syndrome,  
XX CC a chronic inflammatory demyelinating polynuropathy, a hepatobiliary  
XX CC disease, infectious or autoimmune chronic active hepatitis, primary  
XX CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
XX CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
XX CC disease, an autoimmune or immune-mediated skin disease, a bullous skin  
XX CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  
XX CC disease, asthma, allergic rhinitis, atopic dermatitis, food  
XX CC hypersensitivity, urticaria, an immunologic disease of the lung,  
XX CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
XX CC pneumonitis, a transplantation associated disease, graft rejection or  
XX CC graft-versus-host disease. The present sequence represents a PRO protein  
XX CC of the invention.  
XX  
XX SQ Sequence 295 AA;

Query Match 30.8%; Score 472; DB 8; Length 295;  
Best Local Similarity 35.8%; Pred. No. 2e-40;

	Matches	100;	Conservative	52;	Mismatches	101;	Indels	26;	Gaps	4;																																														
Qy	24	GVR	LP	PP	CF	CR	GM	IE	AN	FP	RP	SD	VI	VP	YK	PS	GT	SL	LQ	EV	YV	LV	VS	QG	AD	PE	I	GL	M	N	I	D	83																							
Db	17	GV	PL	IK	YF	AE	AL	GP	LQ	FS	QA	RP	DD	LL	IN	TY	PK	SG	TT	WV	SQ	IL	DM	YI	QG	DL	E	K	C	N	R	A	P	Y	76																					
Qy	84	EOL	PV	LE	V	PQ	P	---	GL	D	I	IK	EL	S	PL	RI	K	SH	LP	VR	FL	PS	DL	NG	DS	K	V	I	Y	M	A	R	N	P	K	D	L	V	139																	
Db	77	VR	PE	LV	ND	P	GS	SG	U	ET	L	K	D	T	P	P	RI	K	SH	L	P	U	L	P	Q	T	LL	D	K	V	V	I	V	A	R	N	P	K	D	V	136															
Qy	140	V	S	Y	Q	F	H	R	S	L	T	M	S	Y	R	G	T	F	Q	B	F	C	R	R	F	M	N	D	K	L	G	Y	S	W	F	H	E	H	R	M	D	S	N	V	L	F	L	K	Y	E	D	M	199			
Db	137	V	S	Y	H	F	H	R	M	E	K	A	H	P	E	G	T	W	S	D	S	F	L	E	K	F	M	A	G	E	V	S	Y	S	W	Y	Q	H	V	O	E	M	S	L	R	T	H	P	V	L	F	Y	E	D	M	196
Qy	200	H	R	D	L	V	T	W	V	E	Q	L	A	R	L	I	G	V	S	C	D	K	A	E	A	L	T	E	H	C	-----	H	Q	L	V	D	O	C	C	N	A	E	A	242												
Db	197	K	E	N	P	K	R	E	I	Q	K	L	E	F	V	R	G	L	P	E	T	W	D	F	M	Q	H	T	S	F	K	E	M	K	N	P	M	T	N	Y	T	T	V	P	Q	E	L	M	D	H	S	I	S	-----	253	
Qy	243	L	P	-	V	G	R	V	G	L	W	K	O	I	F	T	V	S	M	N	E	K	F	D	L	V	Y	K	Q	M	G	K	C	D	L	T	F	280																		
Db	254	-	P	F	W	R	K	G	M	A	G	O	K	W	K	T	T	F	V	A	O	N	E	R	F	P	A	D	A	E	K	M	A	G	C	S	L	F	291																	

RESULT 17	
ADW71803	
ID	ADW71803 standard; protein; 295 AA.
XX	
AC	ADW71803;
XX	
DT	07-APR-2005 (first entry)
XX	
XX	Human phenol sulfotransferase protein, SULT1A3, SEQ ID 4.
XX	
KW	Selectable marker; screening; animal breeding; polymorphism;
KW	restriction fragment length polymorphism; allelic variation;
KW	aryl sulfotransferase; PST; phenol sulfotransferase; SULT1A3; enzyme.
XX	
OS	Homo sapiens.
XX	
XX	Key
XX	Location/Qualifiers
FT	Active-site
FT	83
FT	Active-site
FT	134
FT	Active-site
FT	263
XX	
XX	US2005019788-A1.
XX	
XX	27-JAN-2005.
XX	
XX	30-JAN-2004; 2004US-00769507.
XX	
XX	08-APR-1998; 98US-0081037P.
PR	08-APR-1999; 99US-00288037.
PR	23-NOV-2001; 2001US-00024628.
XX	
XX	(UYGU-) UNIV GUELPH.
PA	
XX	
PI	Squires EJ, Lin Z, Lou Y;
XX	
XX	WPI; 2005-1111909/12.
DR	
XX	Genetically typing animals to determine those with desired boar taint
PT	characteristics, comprises obtaining a sample of genetic material from
PT	the animal, and assaying for the presence of a sulfotransferase allele.
XX	
XX	Example; SEQ ID NO 4; 24pp; English.
PS	
XX	
CC	The present invention relates to a method of genetically typing animals
CC	to determine those with desired boar taint characteristics. The method
CC	involves obtaining a sample of genetic material from the animal and
CC	assaying for the presence of a sulfotransferase allele. The present
CC	sequence is the human phenol sulfotransferase (PST) protein, SULT1A3. The
CC	SULT1A3 protein encoding gene is located on chromosome 16p12.1.
XX	

Query Match 30.8%; Score 472; DB 9; Length 295;  
Best Local Similarity 35.88; Pred. No. 2e-40;  
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;

Qy 24 GVRLPPRCRGMREIANFPVRPSDWMIVTPKSGTSLLOEVLVLSQADPDELGLMNIID 83  
||| : : : ||| : : : ||| : : : ||| : : : ||| : : :  
Db 17 GVPLIKYFABALGLOSFOARPDLLINTYPKSGTTWVSQILDMIYQGDLCKNRAPIY 76  
||| : : : ||| : : : ||| : : : ||| : : : ||| : : :  
Qy 84 EQLPVLEVPPOP---GLDIIEKLTSPLRIKSHLPYRFLPSDLHNGDSKVITYMARNPXDLV 139  
||| : : : ||| : : : ||| : : : ||| : : : ||| : : :  
Db 77 VRVPFLVNDPGPSGLETKIDTPPPRLIIKSHLPLLALPOTLLDQVKVVYVAENPKDVA 136  
||| : : : ||| : : : ||| : : : ||| : : : ||| : : :  
Qy 140 VSYQQFHRSRLTWSYRGTFQFCRRFMNDKLGYSWFHFVQEWRMDSNVLFKYEDM 199  
||| : : : ||| : : : ||| : : : ||| : : : ||| : : :  
Db 137 VSYVHFHRMEKAHPPECTWDISFLEKFWAGEVSYGSWYQHVEWELSRTHPVLVLYFYEDM 196  
||| : : : ||| : : : ||| : : : ||| : : : ||| : : :  
Qy 200 HRDLVTWVEQLARPLGVSCDKAQLEALTEHC-----HQLVDOCCNAEA 242  
||| : : : ||| : : : ||| : : : ||| : : : ||| : : :  
Db 197 KENPKREIQKILEVFGRSLPEETMDFMVQHTSFEMKKKNPMNTYTPVQELMDHSIS--- 253  
||| : : : ||| : : : ||| : : : ||| : : : ||| : : :  
Qy 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKOKMGKCDLTF 280  
||| : : : ||| : : : ||| : : : ||| : : : ||| : : :  
Db 254 -PFMRKGAGDWKTFTTVAQNERFDADYAEKMAGCSLSF 291  
||| : : : ||| : : : ||| : : : ||| : : : ||| : : :

RESULT 18

ID ADI17130

XX AC ADI17130;

XX AC ADI17130;

DT DT

XX 15-APR-2004 (first entry)

DE DE

XX Rat NOVX protein homologue SeqID 666.

KW rat; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;  
KW inflammation; autoimmune disorder; allergy; blood disorder;  
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;  
KW immunoglobulin (IgA) nephropathy; cirrhosis; arthritis;  
KW Alzheimer's disease; infection; str.  
XX XX

OS Rattus norvegicus.

XX OS

PN WO200268649-A2.

XX PN

XX PD 06-SEP-2002.

XX PD

XX PF 31-JAN-2002; 2002WO-US002785.

XX PF

PR 31-JAN-2001; 2001US-0265395P.

PR 31-JAN-2001; 2001US-0265412P.

PR 31-JAN-2001; 2001US-0265514P.

PR 31-JAN-2001; 2001US-0265517P.

PR 02-FEB-2001; 2001US-0266496P.

PR 05-FEB-2001; 2001US-0266767P.

PR 07-FEB-2001; 2001US-0266975P.

PR 07-FEB-2001; 2001US-0267057P.

PR 09-FEB-2001; 2001US-0267459P.

PR 09-FEB-2001; 2001US-0267823P.

PR 15-FEB-2001; 2001US-0268974P.

PR 26-FEB-2001; 2001US-0271664P.

PR 27-FEB-2001; 2001US-0271839P.

PR 27-FEB-2001; 2001US-0271855P.

PR 02-MAR-2001; 2001US-0272788P.

PR 02-MAR-2001; 2001US-0273046P.

PR 14-MAR-2001; 2001US-0275925P.

PR 14-MAR-2001; 2001US-0275947P.

PR 14-MAR-2001; 2001US-0275950P.

PR 14-MAR-2001; 2001US-0275989P.

PR 15-MAR-2001; 2001US-0276448P.

PR 15-MAR-2001; 2001US-0276450P.



```
CC kDa and pI 6.4.
XX Sequence 285 AA;
SQ Query Match      29.9%; Score 458; DB 8; Length 285;
   Best Local Similarity 34.9%; Pred. No. 5.7e-39;
   Matches 96; Conservative 57; Mismatches 104; Indels 18; Gaps 4;

QY 24 GVRLLPFCRGKMEETANFVRPSDVWIVTPKSGTSLQEVVYLVSQADPEIGLMNID 83
DB 7 GIPLIKYFAETIGPLQNFTAWPDDLLISTYPKSGTTWMSEILDMYQGGKLEKCGRAPIY 66
QY 84 EQLPVEYQPQ-----GLDIKELTSPRIKSHLPVRFPLPSDLHNGDSKVIYMARPKDLV 139
DB 67 ARVPLEFKCPVPSGLTLEETAPRLKTHLPUSLPSQLDQKQVVIYIARNAKDV 126
QY 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGVQFWEHRMDSNVFLFKYEDM 199
DB 127 VSYNFYFNMAKLHPDPDGTWDSFLENFMDGEVSGSWYQHKVWELRHTHPVLYLFYEDI 186
QY 200 HRDLVTWVEQLARFLGVSCDKAQLAEALTEHC--HQLVDQC-CNAEALP-----V 245
DB 187 KENPKREIKKLEFLGRSLPEETVDSIVHHTSFKKMKENCMTNTYTIPTIEMDHNVSFPM 246
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
DB 247 RKGTGDMKNTFTVAQNERFDAHYAKTMTDCDFKF 281

RESULT 20
ADD45448
ID ADD45448 standard; protein; 291 AA.
XX
XX ADD45448;
AC
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein P17988, SEQ ID NO 10881.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN WO2003016475-A2.
XX
XX PD
XX PF
XX PF 14-AUG-2002; 2002WO-US025765.
XX
XX PR 14-AUG-2001; 2001US-0312147P.
XX
XX PR 01-NOV-2001; 2001US-0346382P.
XX
XX PR 26-NOV-2001; 2001US-0333347P.
XX
XX PA (GEO) GEN HOSPITAL CORP.
XX
XX PA (FARB) BAYER AG.
XX
XX PI
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
XX
XX DR GENBANK; P17988.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX PS
XX PS Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
```

```
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 291 AA;
```

```
Query Match      29.9%; Score 458; DB 7; Length 291;
Best Local Similarity 34.9%; Pred. No. 5.8e-39;
Matches 96; Conservative 57; Mismatches 104; Indels 18; Gaps 4;

QY 24 GVRLLPFCRGKMEETANFVRPSDVWIVTPKSGTSLQEVVYLVSQADPEIGLMNID 83
DB 13 GIPLIKYFAETIGPLQNFTAWPDDLLISTYPKSGTTWMSEILDMYQGGKLEKCGRAPIY 72
QY 84 EQLPVEYQPQ-----GLDIKELTSPRIKSHLPVRFPLPSDLHNGDSKVIYMARPKDLV 139
DB 73 ARVPLEFKCPVPSGLTLEETAPRLKTHLPUSLPSQLDQKQVVIYIARNAKDV 132
QY 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGVQFWEHRMDSNVFLFKYEDM 199
DB 133 VSYNFYFNMAKLHPDPDGTWDSFLENFMDGEVSGSWYQHKVWELRHTHPVLYLFYEDI 192
QY 200 HRDLVTWVEQLARFLGVSCDKAQLAEALTEHC--HQLVDQC-CNAEALP-----V 245
DB 193 KENPKREIKKLEFLGRSLPEETVDSIVHHTSFKKMKENCMTNTYTIPTIEMDHNVSFPM 252
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
DB 253 RKGTGDMKNTFTVAQNERFDAHYAKTMTDCDFKF 287
```

```
RESULT 21
ADES7139
ID ADES7139 standard; protein; 291 AA.
XX
XX ADES7139;
AC
XX
DT 29-JAN-2004 (first entry)
XX
XX DE Rat Protein P17988, SEQ ID NO 2999.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX OS Rattus norvegicus.
XX
XX PN WO2003016475-A2.
XX
XX PD
XX PD 27-FEB-2003.
XX
XX PF 14-AUG-2002; 2002WO-US025765.
XX
XX PR 14-AUG-2001; 2001US-0312147P.
XX
XX PR 01-NOV-2001; 2001US-0346382P.
XX
XX PR 26-NOV-2001; 2001US-0333347P.
```



```

Db      73  ARVPELEFKCPGPGSGLEETLEETAPRLLKTHLPSSLQSLDQKVYIYARNKDVV 132
Qy      140  VSYQFPHRSLSRTPWSYRGTFQBFECRRPMNDKLGYSWEFHEVQEFWEHRMDSNVLFLEYEDM 199
Db      133  VSYNYFNNAKLLPDPGWDSPLENFMDSGEVSYGTVHQVKWEMLRHTHPVLLEYFYDI 192
Qy      200  HRDLVTWVEQLARFYGVSQCDKAQLEALTEHC--HQLVDQC--CNAEALP-----V 245
Db      193  KENPKREIKIIEFLGRSLPBETVDSIVHHTSFKKMKENCWNTYTTIPTIMDHNVSPEM 252
Qy      246  GRGRVGLWKDIPTVSMNEKFGLVYQKMGKCDLTF 280
Db      253  RKGTTGDMKNTFTVAQNRERFAHAKTMTDCDQPKF 287

```

RESULT 23  
ADE57147  
ID ADE57147 standard; protein; 291 AA.

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that

CC modulates its activity is useful for preparing a medicament for treating  
CC injury (e.g. spinal segmental nerve injury (Chung)), chronic constriction  
CC pain (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 291 AA;

RESULT 24

ADE57143  
ID ADE57143 standard: protein: 291 AA.

XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 291 AA;

Query Match 29.9%; Score 458; DB 7; Length 291;  
 Best Local Similarity 34.9%; Pred. No. 5.8e-39;  
 Matches 96; Conservative 57; Mismatches 104; Indels 18; Gaps 4;

QY 24 GVRLPFCRGKWEETANFPVRPSDVWIVTPKSGTSLQEVVYLVSQGADPDEIGLMNID 83  
 DB 13 GIPLIKYFAETIGPQNTAMPDDLLISTYPKSGTWTWSEIDMTYQGGKLEKCGRAPYI 72  
 QY 84 EQLPVLEYPOP-GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLV 139  
 DB 73 ARVPLEFKCVPVSGLETLTPAPRLKTHLPQLPQLDQKVKVVIARNAKDV 132  
 QY 140 VSYQYFHRSLRTMSYRGTFQFCRRFMDKLGYSWFHVEHFMDSNVFLFKYEDM 199  
 DB 133 VSYNYFNNAKLHPDGTWDSPLENFMDSVSGSWYQHVKEWELRTHPVLVLYFYEDI 192  
 QY 200 HRDLVTWVEQLARFLGVSCDKAQLTEHC--HQLVDQC-CNAEALP-----V 245  
 DB 193 KENPKREIKKLEFLGRSLPETVDSIVHHTSFKKMKNCMTNTYTIPTIEMDHNVSPFM 252  
 QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
 DB 253 RKGTTGDWNTFTVAQNERFDAHYAKTWTDCDFK 287

RESULT 25  
 AAY67294  
 ID AAY67294 standard; protein; 295 AA.  
 XX AAY67294;  
 AC AAY67294;  
 XX 05-APR-2000 (first entry)  
 XX Human STP2 (phenol sulphotransferase 2) amino acid sequence.  
 DE Single nucleotide polymorphism; SNP; STP2; phenol sulphotransferase;  
 XX probe; genotyping; human; drug metabolism.  
 KW Homo sapiens.  
 OS WO9964630-A1.  
 FN  
 XX

PD 16-DEC-1999.  
 XX 09-JUN-1999; 99WO-US013094.  
 XX 10-JUN-1998; 98US-0088710P.  
 XX (AXYS-) AXYS PHARM INC.  
 XX Guida M, Kurth J;  
 XX WPI; 2000-105892/09.  
 DR N-PSDB; AAZ53953.  
 XX Novel nucleic acid used for genotyping, e.g. to predict rate of drug  
 PT metabolism.  
 PT Disclosure; Page 28; 46pp; English.  
 PS This is the human phenol sulphotransferase 2 (STP2) protein sequence. The  
 CC invention relates to sequences AAZ5305-Z59352 which are fragments of the  
 CC STP2 gene. The fragments are from the 8 exons, the promoter region, 3',  
 CC and 5' untranslated regions of the STP2 gene. Each of the sequences  
 CC contains a newly identified STP2 gene single nucleotide polymorphism  
 CC (SNP). STP2 is a phenol sulphotransferase. Substrates for STP2 include  
 CC minoxidil, acetaminophen, and paracetamol. Several of the nucleotide  
 CC changes identified at the polymorphism sites, give rise to an amino acid  
 CC change. Amino acid changes may result in altered enzyme activity. The  
 CC sequences can be used as probes for detecting STP2 polymorphisms. The  
 CC polymorphic probes are used in screening and genotyping, i.e. to predict  
 CC the rate of metabolism of STP2 substrates, potential drug-drug  
 CC interactions and adverse side effects. They can also be used to detect  
 CC diseases resulting from accidental or occupational exposure to toxins and  
 CC to establish animal, cell or in vitro models for drug metabolism  
 XX Sequence 295 AA;

Query Match 29.8%; Score 456; DB 3; Length 295;  
 Best Local Similarity 33.7%; Pred. No. 9.6e-39;  
 Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLPFCRGKWEETANFPVRPSDVWIVTPKSGTSLQEVVYLVSQGADPDEIGLMNID 83  
 DB 17 GVPLIKYFAEALGPLQSFQARPDDLLISTYPKSGTWTWSEIDMTYQGGKLEKCHRAPIF 76  
 QY 84 EQLPVLEYPOP-GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLV 139  
 DB 77 MEVPLEFKVPGISGMETLKNTPAPRLKTHLPQLPQLDQKVKVVIARNAKDV 136  
 QY 140 VSYQYFHRSLRTMSYRGTFQFCRRFMDKLGYSWFHVEHFMDSNVFLFKYEDM 199  
 DB 137 VSYNYFNNAKLHPDGTWDSPLENFMDSVSGSWYQHVKEWELRTHPVLVLYFYEDI 196  
 QY 200 HRDLVTWVEQLARFLGVSCDKAQLTEHC-----HQLVDQCCNAEA 242  
 DB 197 KENPKREIKKLEFLGRSLPETVDSIVHHTSFKKMKNCMTNTYTIPTIEMDHNVSPFM 253  
 QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
 DB 254 -PFMRKGNAGDWKTTFTVAQNERFDAHYAKTWTDCDFK 291

RESULT 26  
 ADP65305  
 ID ADP65305 standard; protein; 295 AA.  
 XX ADP65305;  
 AC ADP65305;  
 XX 12-AUG-2004 (first entry)  
 DT Human sulphotransferase family, cytosolic, 1a.  
 DE autoimmune disease; arthritis; collagen-induced; immunosuppressive; antirheumatic;  
 KW rheumatoid arthritis; collagen-induced; immunosuppressive; antirheumatic;

KW antiarthritic; osteopathic; antigout; antiinflammatory; dermatological;  
KW immunomodulatory; lupus; ankylosing spondylitis; Fibrositis;  
KW fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;  
KW immune; human.  
XX Homo sapiens.  
XX WO2003072827-A1.  
XX PD 04-SEP-2003.  
XX 31-OCT-2002; 2002WO-US035433.  
XX PF 31-OCT-2001; 2001US-0336220P.  
XX PR (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.  
XX PA Hirsch R, Thorton SL;  
XX PI WPI; 2003-712740/67.  
XX DR GENBANK; NP\_001045.  
XX DR Diagnosing and analyzing autoimmune disease using gene expression  
XX PT profiles and microarray technology, useful for diagnosing and treating  
XX PT rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and  
XX PT gout.  
XX PS Disclosure; Page; 56pp; English.  
XX CC The invention relates to a novel method for diagnosing and analysing  
XX CC autoimmune disease or arthritides. The method comprises obtaining a  
XX CC patient sample containing mRNA, analysing gene expression using the mRNA  
XX CC that results in a gene expression signature of the mRNA, and using that  
XX CC gene expression signature to diagnose or analyse the autoimmune disease  
XX CC or arthritides in the patient, where gene expression of at least 60% of  
XX CC the genes correlates with that of the gene signature. The invention  
XX CC further comprises a treatment of rheumatoid arthritis; identification of  
XX CC genes for targeting in the treatment of rheumatoid arthritis in a mammal  
XX CC other than a mouse; diagnosis of rheumatoid arthritis in a mammal; an  
XX CC array or gene chip, specific for rheumatoid arthritis; diagnosis or  
XX CC analyses of autoimmune disease or rheumatoid arthritis; screening the  
XX CC efficacy of a candidate drug in vitro for the treatment of collagen-  
XX CC induced arthritis; and reducing the symptoms associated with collagen-  
XX CC induced arthritis. The compositions of the invention have the following  
XX CC activities: immunosuppressive, antirheumatic, antiarthritic, osteopathic,  
XX CC antigout, antiinflammatory, dermatological, and immunomodulatory. The  
XX CC methods and compositions of the present invention are useful for  
XX CC diagnosing and treating autoimmune disease or arthritides, such as  
XX CC rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis,  
XX CC fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an  
XX CC immune disease caused by an infectious agent. This sequence represents a  
XX CC protein sequence relating to the genes used in the analysis and treatment  
XX CC of autoimmune diseases or arthritides. Note: This sequence is not shown  
XX CC in the specification. It has been supplied in an electronic format from  
XX CC WIPO.

XX SQ Sequence 295 AA;  
Query Match 29.8%; Score 456; DB 7; Length 295;  
Best Local Similarity 33.7%; Pred. No. 9.6e-39;  
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRPPFCRGKMEETANFVPSPDVIVTYPKSGTSLLOEVVYLVSGADPDEIGLMNID 83  
DB 17 GVPLIKYFAEALGPIQSQFARPDLLISTYPSGTTWVSQILDIMYQGGDLEKCHRAPIF 76  
QY 84 EQLPVLVEYQPQ----GLDIIKELTSPRIKSHLPYRFLPSDLHNGDSXVIYMARNPDLV 139  
DB 77 MRVPLEFKVPGIPSGMETLKNTPAPRLKLTHTPLALLPQTLLDQKVYVVARNAKDVA 136  
QY 140 VSYQFHSRLTMSYRGTFQFCRRFMDNKLGYSGWFHVRHMDSNVLFKYEDEM 199  
DB 137 VSYTHFYHMAKYVPHPGTWESGFLEKFMAGEVSYGWYQHVQEWELSRTHPEVLIFYEDM 196

QY 200 HRDLVTWVEOLARFLGVSCDKAQLEALTEHC-----HOLVDQCCNAEA 242  
DB 197 KENPKRIQKILFEVGRSLPEETVDLMVEHTSFEMKKCNPTNTTVRRFMDHSIS--- 253  
QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
DB 254 -PFMRKGMAGDWKTTFTVAQNERFDADYAEKMGACSLSF 291  
RESULT 27  
ADP24022  
ID ADP24022 standard; protein; 295 AA.  
XX AC ADP24022;  
XX DT 18-NOV-2004 (first entry)  
XX DE PRO polypeptide SEQ ID NO:1200.  
XX KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;  
XX KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;  
XX KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.  
XX OS Unidentified.  
XX PN WO2004041170-A2.  
XX PD 21-MAY-2004.  
XX PF 30-OCT-2003; 2003WO-US034312.  
XX PR 01-NOV-2002; 2002US-04233394P.  
XX PA (GETH ) GENENTECH INC.  
XX PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WT;  
XX PI Wu TD;  
XX DR WPI; 2004-419628/39.  
XX DR N-PSDB; ADP24021.  
XX PT New PRO polypeptides and polynucleotides, useful for treating e.g.  
XX PT erythematosis, rheumatoid arthritis, diabetes mellitus, immune-mediated  
XX PT renal disease, or demyelinating diseases of the central or peripheral  
XX PT nervous system.  
XX PS Claim 7; SEQ ID NO 1200; 2940pp; English.  
XX CC The invention relates to a novel isolated nucleic acid and the PRO  
XX CC polypeptide encoded by it. A protein of the invention has  
XX CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,  
XX CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,  
XX CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide  
XX CC of the invention may have a use in gene therapy. The PRO polypeptide, its  
XX CC agonist, antagonist, or antibody that specifically binds to the  
XX CC polypeptide is useful for treating an immune related disorder such as  
XX CC systemic lupus erythematosis, rheumatoid arthritis, osteoarthritis, an  
XX CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
XX CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic  
XX CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
XX CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
XX CC disease, a demyelinating disease of the central or peripheral nervous  
XX CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,  
XX CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary  
XX CC disease, infectious or autoimmune chronic active hepatitis, primary  
XX CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
XX CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
XX CC disease, an autoimmune or immune-mediated skin disease, a bullous skin  
XX CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  
XX CC disease, asthma, allergic rhinitis, atopic dermatitis, food  
XX CC hypersensitivity, urticaria, an immunologic disease of the lung,  
XX CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity



CC pneumonitis, a transplantation associated disease, graft rejection or  
CC graft-versus-host disease. The present sequence represents a PRO protein  
CC of the invention.  
XX  
SQ Sequence 295 AA;  
Query Match 29.8%; Score 456; DB 8; Length 295;  
Best Local Similarity 33.7%; Pred. No. 9.6e-39;  
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;  
QY 24 GVRLPFCRCGKMEETANFVRPSDVWIVTPKSGTSLQEVVLYVSQGADPDEIGLMNID 83  
DB 17 GVPLIKYFAEALGPLOSFOARDLLISTYPSKGTWVSQILDMYQGGDLKCHRAPIF 76  
QY 84 EQLPVEYQPQ----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP KDIV 139  
DB 77 MRVPLEFKVPGIPSGMETLKNTAPARLLKTHPLALLPQTLLDQKVKVYVARNAKDVA 136  
QY 140 VSYIYQFHRSLRTMSYRGTFQFCRRFMDNKLGYGSWFHVEHFMDSNVFLKYEDM 199  
DB 137 VSYIYHFYHMAKVYPHPTGWESLEKFMAGEVSYGSWYQHVEWELSRTHPVLVLYFYEDM 196  
QY 200 HRDLVTMVEQLARFLGVSCDKAQLALTEHC-----HOLVDQCCNAEA 242  
DB 197 KENPKREIQKILFVGRSLPSETVDLMVEHTSFKEKKNQNTYTTVRRFMDHSIS--- 253  
QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
DB 254 -PFMRKMGAGDWKTFTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 28  
ADW71800  
ID ADW71800 standard; protein; 295 AA.  
XX  
AC ADW71800;  
XX  
DT 07-APR-2005 (first entry)  
XX  
DE Pig phenol sulfotransferase protein, SULT1A1, SEQ ID 1.  
XX  
KW Selectable marker; screening; animal breeding; polymorphism;  
KW restriction fragment length polymorphism; allelic variation;  
KW aryl sulfotransferase; PST; phenol sulfotransferase; SULT1A1; enzyme.  
XX  
OS Sus scrofa.  
XX  
FH Key Location/Qualifiers  
FT Active-site 83  
FT Active-site 134  
FT Active-site 263  
XX  
PN US2005019788-A1.  
XX  
PD 27-JAN-2005.  
XX  
PF 30-JAN-2004; 2004US-00769507.  
XX  
PR 08-APR-1998; 98US-0081037P.  
PR 08-APR-1999; 99US-00288037.  
PR 23-NOV-2001; 2001US-00024628.  
XX  
XX (UYGU-) UNIV GUELPH.  
XX  
XX Squires EJ, Lin Z, Lou Y;  
XX  
XX WPI; 2005-111909/12.  
DR N-PSDB; ADW71804.  
XX  
XX Genetically typing animals to determine those with desired boar taint  
XX characteristics, comprises obtaining a sample of genetic material from  
XX the animal, and assaying for the presence of a sulfotransferase allele.

PS Example; SEQ ID NO 1; 24pp; English.  
XX  
CC The present invention relates to a method of genetically typing animals  
CC to determine those with desired boar taint characteristics. The method  
CC involves obtaining a sample of genetic material from the animal and  
CC assaying for the presence of a sulfotransferase allele. The present  
CC sequence is the pig phenol sulfotransferase (PST) protein, SULT1A1.  
XX  
SQ Sequence 295 AA;  
Query Match 29.8%; Score 456; DB 9; Length 295;  
Best Local Similarity 35.6%; Pred. No. 9.6e-39;  
Matches 98; Conservative 51; Mismatches 108; Indels 18; Gaps 3;  
QY 24 GVRLPFCRCGKMEETANFVRPSDVWIVTPKSGTSLQEVVLYVSQGADPDEIGLMNID 83  
DB 17 GVPLIKYFAEALGPLESQAMPDDVLISTYPSKGTWVSEILDLYQGGDLKCHRAPIF 76  
QY 84 EQLPVEYQPQ----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP KDIV 139  
DB 77 MRVPLEFKIPRCPTGFEKLDTPAPRLKTHPLTLLPQTLLDQKVKVYVARNAKDVA 136  
QY 140 VSYIYQFHRSLRTMSYRGTFQFCRRFMDNKLGYGSWFHVEHFMDSNVFLKYEDM 199  
DB 137 VSYIYHFYHMAKVYPNPPTGTFDLMAGEVSYGSWYQHVEWELSRTHPVLVLYFYEDM 196  
QY 200 HRDLVTMVEQLARFLGVSCDKAQLALTEHC---HOLVDQCCNAEALP-----V 245  
DB 197 KENPKREIQKILFVGRSLPSETVEDIVQHTSFQEMKKNAMNTYTLPSDLLDHSISAFM 256  
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
DB 257 RKGITGDKWKTFTTVAQNERFEADYAEKMGAGCNLRF 291  
RESULT 29  
ADW71802  
ID ADW71802 standard; protein; 295 AA.  
XX  
AC ADW71802;  
XX  
DT 07-APR-2005 (first entry)  
XX  
DE Human phenol sulfotransferase protein, SULT1A2, SEQ ID 3.  
XX  
KW Selectable marker; screening; animal breeding; polymorphism;  
KW restriction fragment length polymorphism; allelic variation;  
KW aryl sulfotransferase; PST; phenol sulfotransferase; SULT1A2; enzyme.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Active-site 83  
FT Active-site 134  
FT Active-site 263  
XX  
PN US2005019788-A1.  
XX  
PD 27-JAN-2005.  
XX  
PF 30-JAN-2004; 2004US-00769507.  
XX  
PR 08-APR-1998; 98US-0081037P.  
PR 08-APR-1999; 99US-00288037.  
PR 23-NOV-2001; 2001US-00024628.  
XX  
XX (UYGU-) UNIV GUELPH.  
XX  
XX Squires EJ, Lin Z, Lou Y;  
XX  
XX WPI; 2005-111909/12.  
XX  
XX Genetically typing animals to determine those with desired boar taint

PT characteristics, comprises obtaining a sample of genetic material from  
PT the animal, and assaying for the presence of a sulfotransferase allele.  
XX  
XX  
PS Example; SEQ ID NO 3; 24pp; English.  
XX  
XX The present invention relates to a method of genetically typing animals  
CC to determine those with desired boar taint characteristics. The method  
CC involves obtaining a sample of genetic material from the animal and  
CC assaying for the presence of a sulfotransferase allele. The present  
CC sequence is the human phenol sulfotransferase (PST) protein, SULT1A2. The  
CC SULT1A2 protein encoding gene is located on chromosome 16p12.1.  
XX  
SQ Sequence 295 AA;

Query Match 29.8%; Score 456; DB 9; Length 295;  
Best Local Similarity 33.7%; Pred. No. 9.6e-39;  
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;  
QY 24 GVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLLEQVYLVSQGADPDEIGLMNID 83  
DB 17 GVPLIKYFAEALGPQSQFQAPDDLLISTYPSKGTITWVSQILDMLYQGGDLKCHRAPIF 76  
QY 84 EQLPVLEYQPQ----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLV 139  
DB 77 MRVPLEFKVPCIPSGMETLKNTPAPRLKTHLPLALLPQTLDDQKVKVYVARNAKDVA 136  
QY 140 VSYIYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGVQFWEHRMDSNVLFKYEDM 199  
DB 137 VSYIYHFYHMAKVYPHPTGWESFLEKFMAGEVSYSQYQHVQEWSELSTHFPVLYLFYEDM 196  
QY 200 HRDLVTWVEQLARFLGVSCDKAQLALETHC-----HQLVDQCCNAEA 242  
DB 197 KENPKREIQKILEFVGRSLPBEVDLMVEHTSFKEMKKTPMTNVTYVREFMDHSIS--- 253  
QY 243 LP-VGRGVGLWKDIFTYSMNEKFDLVYKQKMGKCDLTF 280  
DB 254 -PPMRKGMAGDWKTTFTVAQNERPDADYAEKMGAGCSLSF 291

RESULT 30  
ADMW71805  
ID ADMW71805 standard; protein; 295 AA.  
XX  
XX ADMW71805;  
XX  
XX 07-APR-2005 (first entry)  
XX  
XX Pig phenol sulfotransferase protein, SULT1A1, SEQ ID 6.  
XX  
XX Selectable marker; screening; animal breeding; polymorphism;  
KW restriction fragment length polymorphism; allelic variation;  
KW aryl sulfotransferase; PST; phenol sulfotransferase; SULT1A1; enzyme.  
XX  
XX  
XX Sus scrofa.  
XX  
XX Key Location/Qualifiers  
FT Active-site 83  
FT Active-site 134  
FT Active-site 263  
XX  
XX US2005019788-A1.  
XX  
XX 27-JAN-2005.  
XX  
XX 30-JAN-2004; 2004US-00769507.  
XX  
XX 08-APR-1998; 98US-0081037P.  
PR 08-APR-1999; 99US-00288037.  
PR 23-NOV-2001; 2001US-00024628.  
XX  
XX (UYGU-) UNIV GUELPH.  
XX  
XX Squires EJ, Lin Z, Lou Y;

XX WPI; 2005-1111909/12.  
DR N-FSDB; ADMW71804.  
XX  
XX Genetically typing animals to determine those with desired boar taint  
PT characteristics, comprises obtaining a sample of genetic material from  
PT the animal, and assaying for the presence of a sulfotransferase allele.  
XX  
XX Claim 18; SEQ ID NO 6; 24pp; English.  
XX  
XX The present invention relates to a method of genetically typing animals  
CC to determine those with desired boar taint characteristics. The method  
CC involves obtaining a sample of genetic material from the animal and  
CC assaying for the presence of a sulfotransferase allele. The present  
CC sequence is the pig phenol sulfotransferase (PST) protein, SULT1A1.  
XX  
SQ Sequence 295 AA;

Query Match 29.8%; Score 456; DB 9; Length 295;  
Best Local Similarity 35.6%; Pred. No. 9.6e-39;  
Matches 98; Conservative 51; Mismatches 108; Indels 18; Gaps 3;  
QY 24 GVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLLEQVYLVSQGADPDEIGLMNID 83  
DB 17 GVPLIKYFAEALGPLESFQAWPDDVLISTYPSKGTITWVSEILDLYQGGDLKQCORAPIF 76  
QY 84 EQLPVLEYQPQ----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLV 139  
DB 77 VRVPLEFKIPROCTPGFELLKDTAPRLKTHLPLALLPQTLDDQKVKVYVARNAKDVA 136  
QY 140 VSYIYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGVQFWEHRMDSNVLFKYEDM 199  
DB 137 VSYIYHFYHMAKVYPHPTGWESFLEKFMAGEVSYSQYQHVQEWSELSTHFPVLYLFYEDM 196  
QY 200 HRDLVTWVEQLARFLGVSCDKAQLALETHC---HQLVDQCCNAEAALP-----V 245  
DB 197 KENPKREIQKILEFVGRSLPBEETVEDIVQHTSFQEMKNAMNTYRTLPDLLDHSISAFM 256  
QY 246 GRGRVGLWKDIFTYSMNEKFDLVYKQKMGKCDLTF 280  
DB 257 RKGITGDMKSTFTVAQNERFEADYAEKMGAGCNLRF 291

RESULT 31  
ADMW71815  
ID ADMW71815 standard; protein; 295 AA.  
XX  
XX ADMW71815;  
XX  
XX 07-APR-2005 (first entry)  
XX  
XX Pig phenol sulfotransferase mutant protein, SULT1A1.  
XX  
XX Selectable marker; screening; animal breeding; polymorphism;  
KW restriction fragment length polymorphism; allelic variation;  
KW aryl sulfotransferase; PST; phenol sulfotransferase; SULT1A1; enzyme;  
KW mutein.  
XX  
XX Sus scrofa.  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
FT Misc-difference 147  
FT /note= "Wild-type Lys substituted with Glu"  
XX  
XX US2005019788-A1.  
XX  
XX 27-JAN-2005.  
XX  
XX 30-JAN-2004; 2004US-00769507.  
XX  
XX 08-APR-1998; 98US-0081037P.  
PR 08-APR-1999; 99US-00288037.  
PR

PR 23-NOV-2001; 2001US-00024628.  
XX (UYGU-) UNIV GUELPH.  
PA Squires EJ, Lin Z, Lou Y;  
PI WPI: 2005-1111909/12.  
DR N-PSDB; ADW71814.  
XX Genetically typing animals to determine those with desired boar taint  
PT characteristics, comprises obtaining a sample of genetic material from  
PT the animal, and assaying for the presence of a sulfotransferase allele.  
XX Example; Page; 24pp; English.  
XX The present invention relates to a method of genetically typing animals  
CC to determine those with desired boar taint characteristics. The method  
CC involves obtaining a sample of genetic material from the animal and  
CC assaying for the presence of a sulfotransferase allele. The present  
CC sequence is the pig phenol sulfotransferase (PST) mutant protein.  
CC SULT1A1. Note: This sequence is not shown in specification but is derived  
CC from the pig wild-type phenol sulfotransferase shown as SEQ ID NO:6 in  
CC the specification.  
XX Sequence 295 AA;  
SQ  
Query Match 29.7%; Score 454; DB 9; Length 295;  
Best Local Similarity 35.6%; Pred. No. 1.6e-38;  
Matches 98; Conservative 50; Mismatches 109; Indels 18; Gaps 3;  
QY 24 GVRLPFFCRGKMEETANFVRPSDVWIVTPKSGTSLLOEVVYVSQADPDEIGLMNID 83  
DB 17 GVPLIKYFAEALGPLSFQAWDDVLIISTYPKSGTWTWSEILDLYQGGDLQKQCAPIF 76  
QY 84 EQLPVLEYPOP---GLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVIYVARNPKDLV 139  
DB 77 VRVPFLEKIPRCPTGFGELLKDTAPAPRLKTHLPPLTPQLLDQKVVVYVARNAKDVA 136  
QY 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGVQFWEHRMDSNVFLFLKYEDM 199  
DB 137 VSYHYFHYMAEYVYPNGTWSFLEDFMAGEVSGYSQYQHVQSWELRHTHPVLYLFYEDM 196  
QY 200 HRDLVTWVEQLARFLGVSCDKAQLEALTEHC---HQLVDQCCNAEALP-----V 245  
DB 197 KENPKREIQKILEFVGRSLPEETVEDIVQHTSFQEMKQNNAMTYRTLSDLLDHSISAFM 256  
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
DB 257 RKGITGDKWKTFTVAQNERFEADYAEKMGACNLRF 291  
RESULT 32  
ADD18776  
ID ADD18776 standard; protein; 295 AA.  
XX ADD18776;  
XX 15-JAN-2004 (first entry)  
XX Human disease related protein SeqID207.  
XX human; disease state; cytostatic; antiinflammatory; ophthalmological;  
KW antiarteriosclerotic; vulnary; gene therapy;  
KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;  
KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;  
KW glucose transportation; catecholamine synthesis; iron transport;  
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;  
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;  
KW inflammatory condition; wound healing.  
XX Homo sapiens.  
XX WO2003018621-A2.  
PN

XX 06-MAR-2003.  
XX 23-AUG-2002; 2002WO-CB003892.  
XX 23-AUG-2001; 2001GB-00020558.  
PR 05-OCT-2001; 2001GB-00024037.  
XX (OXFO-) OXFORD BIOMEDICA UK LTD.  
XX Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;  
PI WPI: 2003-290046/28.  
DR N-PSDB; ADD18777.  
XX New substantially purified polypeptide, useful for diagnosing or treating  
PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion  
PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or  
PT wound healing.  
XX Claim 25; SEQ ID NO 207; 424pp; English.  
XX This invention relates to novel human genes and gene product which are  
CC implicated in certain disease states. Compounds which modulate the  
CC proteins of the invention may have cytostatic, antiinflammatory,  
CC ophthalmological, antiarteriosclerotic or vulnary activities. The  
CC sequences of the invention may be useful for gene therapy. The invention  
CC may be useful for diagnosing or treating a hypoxia-regulated condition,  
CC such as tumorigenesis, angiogenesis, apoptosis, inflammation,  
CC erythropoiesis, or the biological response to hypoxia conditions  
CC including processes such as glycolysis, gluconeogenesis, glucose  
CC transportation, catecholamine synthesis, iron transport or nitric oxide  
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion  
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,  
CC inflammatory conditions or wound healing. The present sequence is that of  
CC a disease related protein of the invention.  
XX Sequence 295 AA;  
SQ  
Query Match 29.5%; Score 451; DB 7; Length 295;  
Best Local Similarity 33.8%; Pred. No. 3.2e-38;  
Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;  
QY 24 GVRLPFFCRGKMEETANFVRPSDVWIVTPKSGTSLLOEVVYVSQADPDEIGLMNID 83  
DB 17 GVPLIKYFAEALGPLSFQAWDDVLIISTYPKSGTWTWSEILDLYQGGDLQKQCAPIF 76  
QY 84 EQLPVLEYPOP---GLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVIYVARNPKDLV 139  
DB 77 MRVPFLEKAPGIPSGMETLTKDTAPAPRLKTHLPPLTPQLLDQKVVVYVARNAKDVA 136  
QY 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHGVQFWEHRMDSNVFLFLKYEDM 199  
DB 137 VSYHYFHYMAKVPPEPGTWSFLEKFMVGEVSGYSQYQHVQSWELRHTHPVLYLFYEDM 196  
QY 200 HRDLVTWVEQLARFLGVSCDKAQLEALTEHC---HQLVDQCCNAEALP-----V 245  
DB 197 KENPKREIQKILEFVGHSLPEETVDFMVQHTSFKEMKKNPMTNTYTTVPQEFMDHSISPFM 256  
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
DB 257 RKGMDGDKWKTFTVAQNERFDADYAEKMGACLSLF 291  
RESULT 33  
ADE57149  
ID ADE57149 standard; protein; 295 AA.  
XX ADE57149;  
XX 29-JAN-2004 (first entry)  
XX Human Protein P50225, SEQ ID NO 3009.  
DE

```
XX KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX OS Homo sapiens.
XX PN WO2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX PA (GEHO ) GEN HOSPITAL CORP.
XX PA (FARB ) BAYER AG.
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX DR WPI; 2003-268312/26.
XX DR GENBANK; P50225.
XX PT New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX PS Claim 1; Page; 1017pp; English.
XX CC The invention discloses a composition comprising two or more isolated rat
XX CC or human polynucleotides or a polynucleotide which represents a fragment,
XX CC derivative or allelic variation of the nucleic acid sequence. Also
XX CC claimed are a vector comprising the novel polynucleotide, a host cell
XX CC comprising the vector, a method for identifying a nucleotide sequence
XX CC which is differentially regulated in an animal subjected to pain and a
XX CC kit to perform the method, an array, a method for identifying an agent
XX CC that increases or decreases the expression of the polynucleotide sequence
XX CC that is differentially expressed in neuronal tissue of a first animal
XX CC subjected to pain, a method for identifying a compound which regulates
XX CC the expression of a polynucleotide sequence which is differentially
XX CC expressed in an animal subjected to pain, a method for identifying a
XX CC compound that regulates the activity of one or more of the
XX CC polynucleotides, a method for producing a pharmaceutical composition, a
XX CC method for identifying a compound or small molecule that regulates the
XX CC activity in an animal of one or more of the polypeptides given in the
XX CC specification, a method for identifying a compound useful in treating
XX CC pain and a pharmaceutical composition comprising the one or more
XX CC polypeptides or their antibodies. The polynucleotide or the compound that
XX CC modulates its activity is useful for preparing a medicament for treating
XX CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
XX CC therapy). The sequence presented is a human protein (shown in Table 2 of
XX CC the specification) which is differentially expressed during pain. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic form directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 295 AA;

Query Match          29.5%; Score 451; DB 7; Length 295;
Best Local Similarity 33.8%; Pred. No. 3.2e-38;
Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;

QY 24 GVRLLPPFCRGMKEETANFVRPSDVWIVTPYKSGTSLGLEVYLVVSQGDPEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSFQARDDLLISTYPRKSGTTWVSQILDMTYQGDLKCHRPIF 76
QY 84 EQLPVLEYQPQ---GLDIIKELTSPRLIKSHLPYRFLPSSDLHNGDSVIYVARNPDLIV 139
DB 77 MKRVPEFLERKAPIGSGMETKDTAPAPRLKTHLPALLPQTLDDQKVKVYVARNAKDVA 136
QY 140 VSYQYFHRSLRTMSYRGTFQECFRFMNDKLGYSWFEHVFWEHRMDSNVLFILKYEDM 199

Db 137 VSYHYFHYMAKVHPPEPGTWDSFLEKPMVGESVSGWYQHVQWELSKRTHDVLVLYFYEDM 196
QY 200 HRDLVTWVEQLARFLGVSCDKAQALEALTEHC---HQLVDQCCNAEALP-----V 245
Db 197 KENPKREIQKILEFVGRSLPEETVDFMVQHTSFKEMKKNPMTNTYTPVQEFMDHSISPFM 256
QY 246 GRGRVGLWKDIFTYSMNEKFDLVVYKQKMGKCDLTF 280
Db 257 RKGMDAGDKTFTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 34
ADE57137
ID ADE57137 standard; protein; 295 AA.
XX AC ADE57137;
XX DT 29-JAN-2004 (first entry)
XX DE Human Protein P50225, SEQ ID NO 2997.
XX KW Human; pain; neuronal tissue; gene therapy;
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;
XX KW spared nerve injury; SNI; Chung.
XX OS Homo sapiens.
XX PN WO2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX PA (GEHO ) GEN HOSPITAL CORP.
XX PA (FARB ) BAYER AG.
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX DR WPI; 2003-268312/26.
XX DR GENBANK; P50225.
XX PT New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX PS Claim 1; Page; 1017pp; English.
XX CC The invention discloses a composition comprising two or more isolated rat
XX CC or human polynucleotides or a polynucleotide which represents a fragment,
XX CC derivative or allelic variation of the nucleic acid sequence. Also
XX CC claimed are a vector comprising the novel polynucleotide, a host cell
XX CC comprising the vector, a method for identifying a nucleotide sequence
XX CC which is differentially regulated in an animal subjected to pain and a
XX CC kit to perform the method, an array, a method for identifying an agent
XX CC that increases or decreases the expression of the polynucleotide sequence
XX CC that is differentially expressed in neuronal tissue of a first animal
XX CC subjected to pain, a method for identifying a compound which regulates
XX CC the expression of a polynucleotide sequence which is differentially
XX CC expressed in an animal subjected to pain, a method for identifying a
XX CC compound that regulates the activity of one or more of the
XX CC polynucleotides, a method for producing a pharmaceutical composition, a
XX CC method for identifying a compound or small molecule that regulates the
XX CC activity in an animal of one or more of the polypeptides given in the
XX CC specification, a method for identifying a compound useful in treating
XX CC pain and a pharmaceutical composition comprising the one or more
XX CC polypeptides or their antibodies. The polynucleotide or the compound that
XX CC modulates its activity is useful for preparing a medicament for treating
XX CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
XX CC therapy). The sequence presented is a human protein (shown in Table 2 of
XX CC the specification) which is differentially expressed during pain. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic form directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 295 AA;
```

CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 295 AA;

Query Match 29.5%; Score 451; DB 7; Length 295;  
Best Local Similarity 33.8%; Pred. No. 3.2e-38;  
Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;

QY 24 GVRLPPFCRGKMEETANFPVRSDWIVTPKSGTSLQEVVYLVSQGADPDEIGLMNID 83  
DB 17 GVPLIKYFAEALGPLQSFQARPDDLSTPKSGTWTWSQILDMIYQGGDLKCHRAPIF 76  
QY 84 EQLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIVYMARNPOLV 139  
DB 77 MRVPLEFKAIPGIPSGMETLKTDPAPRLKTHPLALLPQTLDDQKVYVYVARNAKDVA 136  
QY 140 VSYIYQFHRSLRTMSYRGTFQEFRCRPMNDKLGYSWFHVFQEFWEHRMDSNVFLFKYEDM 199  
DB 137 VSYHYFYHMAKVHPPEGTWDSFLEKFMVGEVSGSWYQHVQEWELSRTHPVLYLYFYEDM 196  
QY 200 HRDLVTWVQLARFLGVSCDKAQLAETHC---HOLVQCCNAEALP-----V 245  
DB 197 KENPKREIQKILEFVGRSLPEETVDFPMVQHTSFKEKMKKNPMTNITVTPQEFMDHSISPPM 256  
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
DB 257 RKGMAQDWKTTFTVAQNERFDADYAEKMWAGCSLSF 291

RESULT 35  
ADE57141  
ID ADE57141 standard; protein; 295 AA.  
XX  
AC ADE57141;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Human Protein P50225, SEQ ID NO 3001.  
XX  
KW Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
XX  
OS Homo sapiens.  
XX  
PN WO2003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
XX (GHEO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
XX Woolf C, D'urso D, Befort K, Costigan M;  
XX  
XX WPI; 2003-268312/26.  
DR GENBANK; P50225.  
XX  
XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
XX Claim 1; Page; 1017pp; English.  
PS  
XX The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 295 AA;

Query Match 29.5%; Score 451; DB 7; Length 295;  
Best Local Similarity 33.8%; Pred. No. 3.2e-38;  
Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;

QY 24 GVRLPPFCRGKMEETANFPVRSDWIVTPKSGTSLQEVVYLVSQGADPDEIGLMNID 83  
DB 17 GVPLIKYFAEALGPLQSFQARPDDLSTPKSGTWTWSQILDMIYQGGDLKCHRAPIF 76  
QY 84 EQLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIVYMARNPOLV 139  
DB 77 MRVPLEFKAIPGIPSGMETLKTDPAPRLKTHPLALLPQTLDDQKVYVYVARNAKDVA 136  
QY 140 VSYIYQFHRSLRTMSYRGTFQEFRCRPMNDKLGYSWFHVFQEFWEHRMDSNVFLFKYEDM 199  
DB 137 VSYHYFYHMAKVHPPEGTWDSFLEKFMVGEVSGSWYQHVQEWELSRTHPVLYLYFYEDM 196  
QY 200 HRDLVTWVQLARFLGVSCDKAQLAETHC---HOLVQCCNAEALP-----V 245  
DB 197 KENPKREIQKILEFVGRSLPEETVDFPMVQHTSFKEKMKKNPMTNITVTPQEFMDHSISPPM 256  
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
DB 257 RKGMAQDWKTTFTVAQNERFDADYAEKMWAGCSLSF 291

RESULT 36  
ADE57145  
ID ADE57145 standard; protein; 295 AA.  
XX  
AC ADE57145;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Human Protein P50225, SEQ ID NO 3005.  
XX  
KW Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
XX  
OS Homo sapiens.  
XX  
PN WO2003016475-A2.  
XX  
PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.  
 XX PR 14-AUG-2001; 2001US-0312147P.  
 XX PR 01-NOV-2001; 2001US-0346382P.  
 XX PR 26-NOV-2001; 2001US-0333347P.  
 XX PA (GEHO ) GEN HOSPITAL CORP.  
 XX PA (FARB ) BAYER AG.  
 XX PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX WPI; 2003-268312/26.  
 XX DR GENBANK; P50225.  
 XX DR New composition comprising two or more isolated polypeptides, useful for  
 XX PT preparing a medicament for treating pain in an animal.  
 XX PS Claim 1; Page; 1017pp; English.  
 XX CC The invention discloses a composition comprising two or more isolated rat  
 XX CC or human polynucleotides or a polynucleotide which represents a fragment,  
 XX CC derivative or allelic variation of the nucleic acid sequence. Also  
 XX CC claimed are a vector comprising the novel polynucleotide, a host cell  
 XX CC comprising the vector, a method for identifying a nucleotide sequence  
 XX CC which is differentially regulated in an animal subjected to pain and a  
 XX CC kit to perform the method, an array, a method for identifying an agent  
 XX CC that increases or decreases the expression of the polynucleotide sequence  
 XX CC that is differentially expressed in neuronal tissue of a first animal  
 XX CC subjected to pain, a method for identifying a compound which regularly  
 XX CC the expression of a polynucleotide sequence which is differentially  
 XX CC expressed in an animal subjected to pain, a method for identifying a  
 XX CC compound that regulates the activity of one or more of the  
 XX CC polynucleotides, a method for producing a pharmaceutical composition, a  
 XX CC method for identifying a compound or small molecule that regulates the  
 XX CC activity in an animal of one or more of the polypeptides given in the  
 XX CC specification, a method for identifying a compound useful in treating  
 XX CC pain and a pharmaceutical composition comprising the one or more  
 XX CC polypeptides or their antibodies. The polynucleotide or the compound that  
 XX CC modulates its activity is useful for preparing a medicament for treating  
 XX CC pain (e.g. spinal segmental nerve injury (SN1)) in an animal (e.g. gene  
 XX CC injury (CCI) and spared nerve injury (SNI)) in a human protein (shown in Table 2 of  
 XX CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 XX CC the specification) which is differentially expressed during pain. Note:  
 XX CC The sequence data for this patent did not form part of the printed  
 XX CC specification, but was obtained in electronic form directly from WIPO at  
 XX CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX SQ Sequence 295 AA;

Query Match 29.5%; Score 451; DB 7; Length 295;  
 Best Local Similarity 33.8%; Pred. No. 3.2e-38;  
 Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;

QY 24 GVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSLLOEVVYLVSQGADPDEIGLMNID 83  
 DB 17 GVPLIKYFAEALGPLQSFQARPDLLISTYPKSGTWTWSQILDMTYQGGDLKCHRAPIF 76  
 QY 84 EQLPVLVEYQPQ---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDLV 139  
 DB 77 MRVPLEFKAFOIPSGMETLKDTPTAPRLKTHLPALLPQTLLDOKVYVYVARNAKDVA 136  
 QY 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVFQWBFHWRMDSNVLFKLYEDM 199  
 DB 137 VSYHYFHYHMAKVHPPEGTWDSFLEKFMVGEVSGSWYQHVQBWELSRTHPVLVLYFYEDM 196  
 QY 200 HRDLVTWVEQLARFLGVSCDKAQLALETGHC---HQLVDQCCNAALP-----V 245  
 DB 197 KENPKREIQKILEFVGRSLPFEETVDFMVQHTSFKEKMKKNPMTNTYTVQEFMDHSISPPM 256  
 QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
 DB 257 RKGMAGDWKTFTTVAQNERFDADYAEKMGCSLSF 291

## RESULT 37

AD115901  
 ID AD115901 standard; protein; 295 AA.

XX AC AD115901;

XX DT 22-APR-2004 (first entry)

XX DE Human PP 88.

XX KW tumour; antigen; human leukocyte antigen; HLA-A2; HLA-A26;  
 KW cytotoxic T-cell; CTL; vaccine; cancer; colorectal cancer;  
 KW stomach cancer; buccal cancer; renal cancer; lung cancer;  
 KW gynecological cancer; prostate cancer.

XX OS Homo sapiens.

XX PN WO2003008450-A1.

XX PD 30-JAN-2003.

XX PF 11-JUN-2002; 2002WO-JP005799.

XX PR 12-JUN-2001; 2001JP-00177058.

XX PR 21-AUG-2001; 2001JP-00250728.

XX PA (ITOH/) ITOH K.

XX PI Itoh K, Shichijo S;

XX DR WPI; 2003-267996/26.

XX DR N-PSDB; AD115976.

XX PT Tumor antigen peptides recognized by human leukocyte antigen (HLA)-A2 or  
 PT HLA-A26 restricted cytotoxic T-cells for treatment and prevention of  
 PT cancer including preparation of cancer vaccines.

XX PS Claim 2; SEQ ID NO 245; 323pp; Japanese.

XX CC The invention relates to a tumour antigen peptide recognised by human  
 CC leukocyte antigen (HLA)-A2 or HLA-A26 restricted cytotoxic T-cells (CTL)  
 CC and/or capable of inducing CTL. The tumour antigen peptide is useful for  
 CC the treatment, prevention, diagnosis and vaccine production for cancers  
 CC including colorectal, stomach, buccal, renal, lung, gynecological and  
 CC prostate cancer. The present sequence represents the amino acid sequence  
 CC of a human protein.

XX SQ Sequence 295 AA;

Query Match 29.5%; Score 451; DB 7; Length 295;

Best Local Similarity 33.8%; Pred. No. 3.2e-38;

Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;

QY 24 GVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSLLOEVVYLVSQGADPDEIGLMNID 83

DB 17 GVPLIKYFAEALGPLQSFQARPDLLISTYPKSGTWTWSQILDMTYQGGDLKCHRAPIF 76

QY 84 EQLPVLVEYQPQ---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDLV 139

DB 77 MRVPLEFKAFOIPSGMETLKDTPTAPRLKTHLPALLPQTLLDOKVYVYVARNAKDVA 136

QY 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVFQWBFHWRMDSNVLFKLYEDM 199

DB 137 VSYHYFHYHMAKVHPPEGTWDSFLEKFMVGEVSGSWYQHVQBWELSRTHPVLVLYFYEDM 196

QY 200 HRDLVTWVEQLARFLGVSCDKAQLALETGHC---HQLVDQCCNAALP-----V 245

DB 197 KENPKREIQKILEFVGRSLPFEETVDFMVQHTSFKEKMKKNPMTNTYTVQEFMDHSISPPM 256

QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280



CC methods are useful for diagnosing and treating diseases or conditions  
 CC associated with abnormal apoptosis in mammalian tissue, such as cancer,  
 CC inflammation, autoimmune or neurodegenerative disorders. Accordingly,  
 CC they exhibit cytostatic, antiinflammatory, immunosuppressive and  
 CC neuroprotective activities. These may also be used for drug screening  
 CC purposes and in gene therapy. This polypeptide sequence is a human target  
 CC protein, an enzyme associated with the regulation of apoptosis whose  
 CC expression is modulated by novel agents of the invention.

XX Sequence 295 AA;

Query Match 29.5%; Score 451; DB 8; Length 295;  
 Best Local Similarity 33.8%; Pred. No. 3.2e-38;  
 Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;  
 QY 24 GVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLLOEVVYLVSQGADPDEIGLMNID 83  
 DB 17 GVPLIKYFAEALGPLQSFQARPDDLLISTYPKSGTWTWSQILDMIYQGGDLEKCHRAPIF 76  
 QY 84 EQLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDVA 139  
 DB 77 MRVPLEPKAPGPSGMEETLKDTPAPRLKTHLPALLPQTLLDQKVKVYVARNAKDVA 136  
 QY 140 VSYVQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVEHGFHWRMDSNVFLFLKYEDM 199  
 DB 137 VSYVHFYHMAKVHPPEPGTWSFLEKFMVGEVSGSWYQHVQEWELSRTHPVLYLYFYEDM 196  
 QY 200 HRDLVTMVEQLARFLGVSCDKAQLALTEHC---HQLVDQCCNAEALP-----V 245  
 DB 197 KENPKREIQKILEFVGHSLPEETVDFMVQHTSFKEMKKNPMTNTYTPVQEFMDHSISPFM 256  
 QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
 DB 257 RKGWAGDWKTTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 40

ADW71801  
 ID ADW71801 standard; protein; 295 AA.

XX AC ADW71801;

XX DT 07-APR-2005 (first entry)

XX DE Human phenol sulfotransferase protein, SULT1A1, SEQ ID 2.

XX KW Selectable marker; screening; animal breeding; polymorphism;  
 KW restriction fragment length polymorphism; allelic variation;  
 KW aryl sulfotransferase; PST; phenol sulfotransferase; SULT1A1; enzyme.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Active-site 83

FT Active-site 134

FT Active-site 263

XX US2005019788-A1.

XX PD 27-JAN-2005.

XX PF 30-JAN-2004; 2004US-00769507.

XX PR 08-APR-1998; 98US-0081037P.

XX PR 08-APR-1999; 99US-00288037.

XX PR 23-NOV-2001; 2001US-00024628.

XX PA (UYGU-) UNIV GUELPH.

XX PI Squires EJ, Lin Z, Lou Y;

XX DR WPI; 2005-111909/12.

XX

PT Genetically typing animals to determine those with desired boar taint  
 PT characteristics, comprises obtaining a sample of genetic material from  
 PT the animal, and assaying for the presence of a sulfotransferase allele.  
 XX Example; SEQ ID NO 2; 24pp; English.

XX The present invention relates to a method of genetically typing animals  
 CC to determine those with desired boar taint characteristics. The method  
 CC involves obtaining a sample of genetic material from the animal and  
 CC assaying for the presence of a sulfotransferase allele. The present  
 CC sequence is the human phenol sulfotransferase (PST) protein, SULT1A1. The  
 CC SULT1A1 protein encoding gene is located on chromosome 16p12.1.

SQ Sequence 295 AA;

Query Match 29.5%; Score 451; DB 9; Length 295;  
 Best Local Similarity 33.8%; Pred. No. 3.2e-38;  
 Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;  
 QY 24 GVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLLOEVVYLVSQGADPDEIGLMNID 83  
 DB 17 GVPLIKYFAEALGPLQSFQARPDDLLISTYPKSGTWTWSQILDMIYQGGDLEKCHRAPIF 76  
 QY 84 EQLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDVA 139  
 DB 77 MRVPLEPKAPGPSGMEETLKDTPAPRLKTHLPALLPQTLLDQKVKVYVARNAKDVA 136  
 QY 140 VSYVQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVEHGFHWRMDSNVFLFLKYEDM 199  
 DB 137 VSYVHFYHMAKVHPPEPGTWSFLEKFMVGEVSGSWYQHVQEWELSRTHPVLYLYFYEDM 196  
 QY 200 HRDLVTMVEQLARFLGVSCDKAQLALTEHC---HQLVDQCCNAEALP-----V 245  
 DB 197 KENPKREIQKILEFVGHSLPEETVDFMVQHTSFKEMKKNPMTNTYTPVQEFMDHSISPFM 256  
 QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
 DB 257 RKGWAGDWKTTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 41

AD117050

ID AD117050 standard; protein; 302 AA.

XX AC AD117050;

XX DT 15-APR-2004 (first entry)

XX DE Human NOVX protein homologue SeqID 586.

XX KW human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;  
 KW inflammation; autoimmune disorder; allergy; blood disorder;  
 KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;  
 KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;  
 KW Alzheimer's disease; infection; str.

XX OS Homo sapiens.

XX PN WO200268649-A2.

XX PD 06-SEP-2002.

XX PF 31-JAN-2002; 2002WO-US0002785.

XX PR 31-JAN-2001; 2001US-0285395P.

XX PR 31-JAN-2001; 2001US-0265412P.

XX PR 31-JAN-2001; 2001US-0265514P.

XX PR 31-JAN-2001; 2001US-0265517P.

XX PR 02-FEB-2001; 2001US-0266406P.

XX PR 05-FEB-2001; 2001US-0266767P.

XX PR 07-FEB-2001; 2001US-0266975P.

XX PR 07-FEB-2001; 2001US-0267057P.

XX PR 08-FEB-2001; 2001US-0267459P.



09-FEB-2001; 2001US-0267823P.  
15-FEB-2001; 2001US-0268974P.  
26-FEB-2001; 2001US-0271664P.  
27-FEB-2001; 2001US-0271839P.  
27-FEB-2001; 2001US-0271855P.  
02-MAR-2001; 2001US-0272788P.  
02-MAR-2001; 2001US-0273046P.  
14-MAR-2001; 2001US-0275925P.  
14-MAR-2001; 2001US-0275947P.  
14-MAR-2001; 2001US-0275950P.  
15-MAR-2001; 2001US-0275989P.  
15-MAR-2001; 2001US-0276448P.  
15-MAR-2001; 2001US-0276450P.  
16-MAR-2001; 2001US-0276397P.  
16-MAR-2001; 2001US-0276768P.  
20-MAR-2001; 2001US-0278652P.  
26-MAR-2001; 2001US-0278775P.  
26-MAR-2001; 2001US-0278778P.  
29-MAR-2001; 2001US-0279882P.  
29-MAR-2001; 2001US-0279884P.  
30-MAR-2001; 2001US-0280147P.  
11-APR-2001; 2001US-0282992P.  
11-APR-2001; 2001US-0283083P.  
20-APR-2001; 2001US-0285133P.  
23-APR-2001; 2001US-0285749P.  
03-MAY-2001; 2001US-0288327P.  
03-MAY-2001; 2001US-0288504P.  
29-MAY-2001; 2001US-0294047P.  
30-MAY-2001; 2001US-0294473P.  
08-JUN-2001; 2001US-0296964P.  
18-JUN-2001; 2001US-0298959P.  
19-JUN-2001; 2001US-0299324P.  
13-AUG-2001; 2001US-0312020P.  
16-AUG-2001; 2001US-0312889P.  
21-AUG-2001; 2001US-0312908P.  
28-AUG-2001; 2001US-0313390P.  
31-AUG-2001; 2001US-0315470P.  
07-SEP-2001; 2001US-0316447P.  
07-SEP-2001; 2001US-0318115P.  
12-SEP-2001; 2001US-0318118P.  
19-SEP-2001; 2001US-0318740P.  
18-OCT-2001; 2001US-0323379P.  
18-OCT-2001; 2001US-0330245P.  
14-NOV-2001; 2001US-0332701P.  
(CURA-) CURAGEN CORP.  
Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;  
Li L, Gangolli EA, Padigar M, Anderson DW, Rastelli L, Miller CE;  
Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CE;  
Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;  
WPI; 2002-706998/76.  
New NOVX polypeptides and nucleic acids, useful for preventing or  
treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,  
atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or  
pharmacogenomics.  
Disclosure; SEQ ID NO 586; 1498pp; English.

CC treating or preventing diseases such as inflammation, autoimmune  
disorders, allergies, blood disorders, acquired immunodeficiency syndrome  
(AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,  
arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy  
and epilepsy. Accordingly, these molecules have many activities including  
cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,  
CC haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,  
CC antiasthmatic, nephrotropic, antiarthritic, hepatotropic,  
CC neuroprotective, nootropic, antibacterial, viricide, antiparasitic,  
CC relaxant and anticonvulsant. In addition, they are useful in screening  
assays to identify small molecules that modulate or inhibit, for example,  
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also  
used as in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein  
of the invention.  
XX  
SQ Sequence 302 AA;  
Query Match 29.3%; Score 449; DB 5; Length 302;  
Best Local Similarity 36.2%; Pred. No. 5.4e-38;  
Matches 96; Conservative 52; Mismatches 99; Indels 18; Gaps 3;  
QY 36 EEIANFPVRPSDVMTVTPKSGTSLQEVVLVSGADPDEIGLMNIDEQLPVLVYQPQ- 94  
Db 36 DKIMNFQAKPDDLLISTYPKAGTTWTQEIVELIQNEGDEKSKRAPHQRPFFLEWKIPS 95  
QY 95 ---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVITYMARNPKDLVVSYQPHSLRT 151  
Db 96 LGSGLAQAHAMPSPRILKTHLPFHLPPSLLEKCNKIIIVARNPKDNMVSYVHFQRMNKA 155  
QY 152 MSYRGTFQFCRRFMNDKLGYSWFHVEHQRMSNVLFKYEDMHRDLVTWVEQLA 211  
Db 156 LPAPGTWEEYFETFLAGKVCWGSWHEHVKGWWEAKDHRILYLFYEDMKCKPKHIEQLA 215  
QY 212 RFLGVSCDKAOLEALTEHCQLV---DOCCNAEALP-----VGRGVGLWKDIF 257  
Db 216 EFIGKLDKVDKLVHYTSFDMVKQNPANYSSIPAEIMDHSISPFMRKGAUGDWKKHF 275  
QY 258 TVSMNEKFDLVYKQMGKCDLTDF 282  
Db 276 TVAQNERFDEYKKQMTDRLTFHF 300  
RESULT 42  
ADF76915  
ID ADF76915 standard; protein; 302 AA.  
XX  
AC ADF76915;  
XX  
DT 26-FEB-2004 (first entry)  
XX  
DE Novel human secreted and transmembrane protein SeqID 590.  
XX  
KW human; PRO; membrane bound protein; membrane bound receptor;  
KW cell proliferation; cell migration; cell differentiation;  
KW mitogenic factor; survival factor; cytotoxic factor;  
KW differentiation factor; neuropeptide; hormone; cell receptor;  
KW receptor-ligand interaction; cytostatic; chondrocyte; tumour.  
XX  
OS Homo sapiens.  
XX  
PN WO2003072035-A2.  
XX  
PD 04-SEP-2003.  
XX  
PF 21-FEB-2003; 2003WO-US005241.  
XX  
PR 22-FEB-2002; 2002US-0359461P.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR;  
PI Williams PM, Wood WI, Wu TD;





```

Db      248 IMDQISPPMRKGI VGDWKNHFTVAQNERFDEIVEQKMDGTSLNF 292

RESULT 45
ID      ABM84002
AB      ABM84002 standard; protein; 327 AA.
XX
XX
AC      ABM84002;
XX
DT      18-NOV-2004 (first entry)
XX
DE      Human diagnostic and therapeutic pprotein SEQ ID NO:4251.
XX
KW      gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
OS      Homo sapiens.
XX
PN      WO2004023973-A2.
XX
PD      25-MAR-2004.
XX
PF      12-SEP-2003; 2003WO-US028227.
PR      12-SEP-2002; 2002US-0410259P.
PR      12-SEP-2002; 2002US-0410260P.
XX
PA      (INCY-) INCYTE CORP.
XX
PI      Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI      Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI      Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI      Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI      Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI      Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI      Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI      Patury S, Shi X, Suarez CJ;
XX
DR      WPI; 2004-329368/30.
DR      N-PSDB; ACN42654.
XX
PT      New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT      in diagnosing a condition, disease or disorder associated with human
PT      molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT      in gene mapping.
XX
PS      Claim 27; Page; 190pp; English.
XX
CC      The invention relates to novel diagnostic and therapeutic polynucleotides
CC      selected from one of the 2722 sequences defined in the specification. A
CC      polynucleotide of the invention may have a use in gene therapy. The human
CC      diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC      used to diagnose a particular condition, disease or disorder associated
CC      with human molecules, e.g. cell proliferative disorders,
CC      autoimmune/inflammatory disorder, developmental disorder, endocrine
CC      disorder, neurological disorders, gastrointestinal disorders, or
CC      infections caused by virus, bacteria, fungi or parasite. The dithp
CC      molecules may also be used in genetic mapping, in identifying individuals
CC      from minute biological samples, in detecting single nucleotide
CC      polymorphisms, as molecular weight markers, and for somatic or germline
CC      gene therapy. The present sequence represents a dithp protein of the
CC      invention. Note: The sequence data for this patent is not represented in
CC      the printed specification, but was obtained in electronic format directly
CC      from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ      Sequence 327 AA;

Query Match      29.2%; Score 447; DB 8; Length 327;
Best Local Similarity 32.5%; Pred. No. 9.8e-38;
Matches 101; Conservative 51; Mismatches 101; Indels 58; Gaps 5;

QY      24 GVRLPFFCRGKMEETANFVRPSDVWIVTYPKSGTSLQEVVYLVSQGADPDEIGLMNID 83
      ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB      17 GVPLIKYFAEALGPLQSFQARPDDLINTYPKSGTWTWSQILDMYQGGDLKCNRAPIY 76

```

CC disorder, neurological disorders, gastrointestinal disorders, or  
CC infections caused by virus, bacteria, fungi or parasite. The dithp  
CC molecules may also be used in genetic mapping, in identifying individuals  
CC from minute biological samples, in detecting single nucleotide  
CC polymorphisms, as molecular weight markers, and for somatic or germline  
CC gene therapy. The present sequence represents a dithp protein of the  
CC invention. Note: The sequence data for this patent is not represented in  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)  
XX  
XX  
SQ Sequence 327 AA;

Query Match 29.2%; Score 447; DB 8; Length 327;  
Best Local Similarity 32.5%; Pred. No. 9.8e-38;  
Matches 101; Conservative 51; Mismatches 101; Indels 58; Gaps 5;  
QY 24 GVRLLPPFCRGKMEIANFVRPSDVWIVTYPKSGTSLLOEVVLYVSQGDADPEIGLMNID 83  
DB 17 GVPLIKYFAEALGPLQSFQARPDLLINTYPSGTTWVSQILDMIYQGGDLKCNRAPIY 76  
QY 84 EQLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP KDIV 139  
DB 77 VRVPFLEVNDPGEPSGLETLDTPPEPLIKSHLPALLPQTLLDQKVKVYVARNPKDVA 136  
QY 140 VSYVYOFHRSRLTMSYRGTFQFCRRFMDKLG----- 171  
DB 137 VSYVYHFRMEKAHPPEPGTWDSEFLKFMAGEGGLDWRKEGKVPGRGGYVNVQPCVGAACP 196  
QY 172 ---YGSWFHVRQFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAOLEALT 227  
DB 197 LXVSGSWYQHVQEWELSRTHPVLYLFYEDMKNPKREIQKILEFVGRSLPEETMDFMV 256  
QY 228 EHC-----HOLVQCCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVY 269  
DB 257 QHTSFKMKKNPMTNVTYVPOELMDHSIS----PFMRKGMAGDKWKTFTVAQNERFDADY 312  
QY 270 KQKMGKCDLTF 280  
DB 313 AEKMGCSLSF 323

## RESULT 47

ABM82615  
ID ABM82615 standard; protein; 327 AA.

XX AC ABM82615;  
XX DT 18-NOV-2004 (first entry)  
XX DE Human diagnostic and therapeutic pprotein SEQ ID NO:2864.  
XX KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.  
XX OS Homo sapiens.

XX WO2004023973-A2.

XX 25-MAR-2004.

XX 12-SEP-2003; 2003WO-US028227.

XX 12-SEP-2002; 2002US-0410259P.

XX 12-SEP-2002; 2002US-0410260P.

XX (INCY-) INCYTE CORP.

XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;  
PI Harthshorne TA, Suchorski MT, Altus CM, Pitts SJ, Elder LV;  
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;  
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;  
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;  
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;  
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;

PI Patury S, Shi X, Suarez CJ;  
XX WPI; 2004-329368/30.  
XX N-PSDB; ACN41267.  
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful  
XX in diagnosing a condition, disease or disorder associated with human  
XX molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or  
XX in gene mapping.  
XX  
XX PS Claim 27; Page; 190pp; English.  
XX  
XX CC The invention relates to novel diagnostic and therapeutic polynucleotides  
XX selected from one of the 2722 sequences defined in the specification. A  
XX polynucleotide of the invention may have a use in gene therapy. The human  
XX diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be  
XX used to diagnose a particular condition, disease or disorder associated  
XX with human molecules, e.g. cell proliferative disorders,  
XX autoimmune/inflammatory disorder, developmental disorder, endocrine  
XX disorder, neurological disorders, gastrointestinal disorders, or  
XX infections caused by virus, bacteria, fungi or parasite. The dithp  
XX molecules may also be used in genetic mapping, in identifying individuals  
XX from minute biological samples, in detecting single nucleotide  
XX polymorphisms, as molecular weight markers, and for somatic or germline  
XX gene therapy. The present sequence represents a dithp protein of the  
XX invention. Note: The sequence data for this patent is not represented in  
XX the printed specification, but was obtained in electronic format directly  
XX from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)  
XX  
XX SQ Sequence 327 AA;

Query Match 29.2%; Score 447; DB 8; Length 327;  
Best Local Similarity 32.5%; Pred. No. 9.8e-38;  
Matches 101; Conservative 51; Mismatches 101; Indels 58; Gaps 5;  
QY 24 GVRLLPPFCRGKMEIANFVRPSDVWIVTYPKSGTSLLOEVVLYVSQGDADPEIGLMNID 83  
DB 17 GVPLIKYFAEALGPLQSFQARPDLLINTYPSGTTWVSQILDMIYQGGDLKCNRAPIY 76  
QY 84 EQLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP KDIV 139  
DB 77 VRVPFLEVNDPGEPSGLETLDTPPEPLIKSHLPALLPQTLLDQKVKVYVARNPKDVA 136  
QY 140 VSYVYOFHRSRLTMSYRGTFQFCRRFMDKLG----- 171  
DB 137 VSYVYHFRMEKAHPPEPGTWDSEFLKFMAGEGGLDWRKEGKVPGRGGYVNVQPCVGAACP 196  
QY 172 ---YGSWFHVRQFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAOLEALT 227  
DB 197 LXVSGSWYQHVQEWELSRTHPVLYLFYEDMKNPKREIQKILEFVGRSLPEETMDFMV 256  
QY 228 EHC-----HOLVQCCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVY 269  
DB 257 QHTSFKMKKNPMTNVTYVPOELMDHSIS----PFMRKGMAGDKWKTFTVAQNERFDADY 312  
QY 270 KQKMGKCDLTF 280  
DB 313 AEKMGCSLSF 323

## RESULT 48

ABM84003

ID ABM84003 standard; protein; 327 AA.

XX AC ABM84003;

XX DT 18-NOV-2004 (first entry)

XX DE Human diagnostic and therapeutic pprotein SEQ ID NO:4252.

XX KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.

XX OS Homo sapiens.

```
XX PN WO2004023973-A2.
XX PD 25-MAR-2004.
XX PF 12-SEP-2003; 2003WO-US028227.
XX PR 12-SEP-2002; 2002US-0410259P.
XX PR 12-SEP-2002; 2002US-0410260P.
XX PA (INCY-) INCYTE CORP.
XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthehorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtson ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.
DR N-PSDB; ACN42655.
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX PS Claim 27; Page; 190pp; English.
XX CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX SQ Sequence 327 AA;
Query Match 29.2%; Score 447; DB 8; Length 327;
Best Local Similarity 32.5%; Pred. No. 9.8e-38;
Matches 101; Conservative 51; Mismatches 101; Indels 58; Gaps 5;
QY 24 GVRLPFCRGKMEETANFVRPSDVWIVTPKSGTSLQEVVYLSQCADPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPQSQARDDLLINTPKSGTWTWSQILDMIYQGGDLKCNRAPIY 76
QY 84 EQLPVLEYQPQ---GLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVIYMARPKDLV 139
DB 77 VRVPFLVNDQPEPSGLETLDTPPRLIKSHLPALPLPQTLDDQKVYVVARPNKQVA 136
QY 140 VSYIYFHRSLRTMSVGTFTQFCRRPMDKLG----- 171
DB 137 VSYIYFHRMEKAHPPEPGTWDSPLEKFXMAGEGLDWRKEGVKPRGGYVNVQPCVGAQPL 196
QY 172 ---YGSNFEHVQEFWEHRMDSNVLFKLYEDMHRDLVTMVEQLARFLGVSCDKQALEALT 227
DB 197 LXVSGSYQHVOEWELSRTHPVLYLFYEDMKENPKREIQKILFVGRSLPEETMDPMV 256
QY 228 EHC-----HQLVDOCCNAEALP-VGRGRVGLMKDIFTVSNEKFDLVY 269
DB 257 QHTSPKMKKNPMTNTYTVPQELMDHSIS----PFMRKMGADGWKTTFVAGNERFDADY 312
```

```
QY 270 KQKMGKCDLTF 280
DB 313 AERWAGCSLSF 323
RESULT 49
ABM84004
ID ABM84004 standard; protein; 327 AA.
XX AC ABM84004;
XX DT 18-NOV-2004 (first entry)
XX DE Human diagnostic and therapeutic pprotein SEQ ID NO:4253.
XX KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX OS Homo sapiens.
XX PN WO2004023973-A2.
XX PD 25-MAR-2004.
XX PF 12-SEP-2003; 2003WO-US028227.
XX PR 12-SEP-2002; 2002US-0410259P.
XX PR 12-SEP-2002; 2002US-0410260P.
XX PA (INCY-) INCYTE CORP.
XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthehorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtson ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.
DR N-PSDB; ACN42656.
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX PS Claim 27; Page; 190pp; English.
XX CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX SQ Sequence 327 AA;
Query Match 29.2%; Score 447; DB 8; Length 327;
Best Local Similarity 32.5%; Pred. No. 9.8e-38;
Matches 101; Conservative 51; Mismatches 101; Indels 58; Gaps 5;
```



```
XX SQ Sequence 269 AA;
Query Match 29.0%; Score 444; DB 5; Length 269;
Best Local Similarity 40.6%; Pred. No. 1.5e-37;
Matches 104; Conservative 39; Mismatches 83; Indels 30; Gaps 7;

Qy 41 FVPSDVIWYTPKSGTSLQEVVLYSQAD-----PDEIGLMNIDEQLPVLEYPPQG 95
Db 20 FOARDDVLIAGYPKSGTWTQLQELLSLHPNVGDFEPSDPLLFNR-----PWLEYPK-G 73

Qy 96 LDIKEL----TSPLIKSHLPYRFLPSDLHNGDSKVITYMARNPKDLVVSYVYQFHSRLRT 151
Db 74 EDWYETLAPMSSPLIKTHPLELLPKSFLSSKAKIIVLRNPKDVAVSYHFSRSHKD 133

Qy 152 M-SYRGTFQECRRPMNDKLGVSFWFHVQBFWEHRMDSNVFLKYEDMHRDLVTMVQOL 210
Db 134 LPADPGTPEEFLEAFNGKVLGYSYFDHVLGWELRPEPQVFLFDYEDLKEDPAGEIKKI 193

Qy 211 ARFLGVSCDKAQLALETCH---QLVDOCCNAEALPVG-----BGRVGLWKDI 256
Db 194 AEFGLPLSEBELDKLLOHSPFLMKJNPLSNYETLCIGKSKGRKSPFMRKGLVGDWKNY 253

Qy 257 FTVSMNEKFDLVYKOK 272
Db 254 FTPEQNEKFDVKEK 269

RESULT 51
AD117331
ID AD117331 standard; protein; 269 AA.
XX AC AD117331;
XX DT 15-APR-2004 (first entry)
XX DE Polypeptide homologous to a human NOVX domain SeqID 867.
XX KW NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes; inflammation;
KW autoimmune disorder; allergy; blood disorder;
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
KW Alzheimer's disease; infection; str.
XX OS Unidentified.
XX PN WO200268649-A2.
XX PD 06-SEP-2002.
XX PF 31-JAN-2002; 2002WO-US002785.
XX PR 31-JAN-2001; 2001US-0265395P.
PR 31-JAN-2001; 2001US-0265412P.
PR 31-JAN-2001; 2001US-0265514P.
PR 31-JAN-2001; 2001US-0265517P.
PR 02-FEB-2001; 2001US-0266408P.
PR 05-FEB-2001; 2001US-0266767P.
PR 07-FEB-2001; 2001US-0266975P.
PR 08-FEB-2001; 2001US-0267057P.
PR 09-FEB-2001; 2001US-0267459P.
PR 15-FEB-2001; 2001US-0267823P.
PR 15-FEB-2001; 2001US-0268974P.
PR 26-FEB-2001; 2001US-0271664P.
PR 27-FEB-2001; 2001US-0271839P.
PR 27-FEB-2001; 2001US-0271855P.
PR 02-MAR-2001; 2001US-0272789P.
PR 02-MAR-2001; 2001US-0273046P.
PR 14-MAR-2001; 2001US-0275925P.
PR 14-MAR-2001; 2001US-0275947P.
PR 14-MAR-2001; 2001US-0275950P.
PR 14-MAR-2001; 2001US-0275989P.
PR 15-MAR-2001; 2001US-0276448P.
PR 15-MAR-2001; 2001US-0276450P.
PR 16-MAR-2001; 2001US-0276397P.
PR 16-MAR-2001; 2001US-0276786P.
PR 20-MAR-2001; 2001US-0278652P.
PR 26-MAR-2001; 2001US-0278775P.
PR 26-MAR-2001; 2001US-0278778P.
PR 29-MAR-2001; 2001US-0279882P.
PR 29-MAR-2001; 2001US-0279884P.
PR 30-MAR-2001; 2001US-0280147P.
PR 11-APR-2001; 2001US-0282992P.
PR 11-APR-2001; 2001US-0283083P.
PR 20-APR-2001; 2001US-0285133P.
PR 23-APR-2001; 2001US-0285749P.
PR 03-MAY-2001; 2001US-0288327P.
PR 03-MAY-2001; 2001US-0288504P.
PR 29-MAY-2001; 2001US-0294047P.
PR 30-MAY-2001; 2001US-0294473P.
PR 08-JUN-2001; 2001US-0296964P.
PR 18-JUN-2001; 2001US-0298959P.
PR 19-JUN-2001; 2001US-0299324P.
PR 13-AUG-2001; 2001US-0312020P.
PR 16-AUG-2001; 2001US-0312889P.
PR 21-AUG-2001; 2001US-0313390P.
PR 28-AUG-2001; 2001US-0315470P.
PR 31-AUG-2001; 2001US-0316447P.
PR 07-SEP-2001; 2001US-0318115P.
PR 07-SEP-2001; 2001US-0318118P.
PR 12-SEP-2001; 2001US-0318740P.
PR 19-SEP-2001; 2001US-0323379P.
PR 18-OCT-2001; 2001US-0330245P.
PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.
XX PA (CURA-) CURAGEN CORP.
XX PT Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;
PI Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;
PI Gerlach VL, Taupier RU, Gusev VY, Colman SD, Wolenc AR, Pena CEA;
PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
XX WPI; 2002-706998/76.
XX PT New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX PS Disclosure; SEQ ID NO 867; 1498pp; English.
XX CC This invention relates to a novel nucleic acids, and encoded polypeptides
CC thereof, which have properties related to the stimulation of biochemical
CC or physiological responses in a cell, tissue, organ or organism.
CC Specifically, it refers to the use of biologically active fragments for
CC diagnostic and prognostic assays and furthermore in the treatment of
CC diverse pathological conditions. The present invention describes novel
CC human and murine NOVX proteins, as well as methods to modulate their
CC expression using antisense oligos, ribozymes and peptide nucleic acids.
CC The NOVX polypeptides, polynucleotides and antibodies are useful in
CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in
CC treating or preventing diseases such as inflammation, autoimmune
CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome
CC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,
CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
CC and epilepsy. Accordingly, these molecules have many activities including
CC haemostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,
CC cyostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
CC antiasthmatic, nephrotropic, antiarthritic, hepatotropic,
CC neuroprotective, neurotropic, antibacterial, virucide, antiparasitic.
CC relaxant and anticonvulsant. In addition, they are useful in screening
CC assays to identify small molecules that modulate or inhibit, for example,
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
```



CC used as in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. This polypeptide is a protein fragment that is  
CC homologous to a domain of a human NOVX protein of the invention.  
XX  
SQ Sequence 269 AA;  
Query Match 29.0%; Score 444; DB 5; Length 269;  
Best Local Similarity 40.6%; Pred. No. 1.5e-37;  
Matches 104; Conservative 39; Mismatches 83; Indels 30; Gaps 7;  
QY 41 FVRRPSDVIVTPKSGTSLLOEVVLYSQGAD-----PDEIGLMNIDQLPVLVYQPG 95  
DB 20 FQARPDVLIAGVPGTTLQELSLHENVGDFEPPSPDPLFRN-----PWLEVPK-G 73  
QY 96 LDIIEKL-----TSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDLVVSYYQFHRSLRT 151  
DB 74 EDWYETLKPMPSPRLIKTHLPLELLPKSFLSSKAKIIVLRNPKDVAVSYHFRSRHKD 133  
QY 152 M-SYRGTFQFCRRFNMNDKLGVSQFHVQEFWEHRMDSNVFLKAYEDMHRDLVTMVEQL 210  
DB 134 LPADPGTFEELEAFNGHVKVLYGSPFDHVLGWELRPEQVFLFDYEDLKEDPAGEIKKI 193  
QY 211 ARFLGVSCDKAOLEALTECH---OLVDOCCNAEALPVG-----RGRVGLWKDI 256  
DB 194 AEFGLPLSEBELDKLHSSFFLMKLNPLSNVETLCGSKGRKSPFMRKGLVGDWKNY 253  
QY 257 FTVSMNEKPDLYVYKOK 272  
DB 254 FTPEQNEKPKVKEK 269  
RESULT 52  
ADD12573  
ID ADD12573 standard; protein; 310 AA.  
XX  
AC ADD12573;  
XX  
DT 01-JAN-2004 (first entry)  
DE Human ENZM-33 protein SEQ ID NO:33.  
XX  
KW human; enzyme; ENZM; cytostatic; antiarteriosclerotic; anti-HIV;  
KW neurologic; cerebroprotective; antiparkinsonian; anticonvulsant;  
KW nontropic; neuroprotective; antiinflammatory; ophthalmological;  
KW antithyroid; antirheumatic; antibacterial; virucide; protozoacide;  
KW antiparasitic; fungicide; anorectic; cardiant; hypotensive;  
KW antinfertility; hepatotropic; gene therapy; autoimmune disorder;  
KW inflammatory disorder; AIDS; allergy; atopic dermatitis; arthritis;  
KW thyroiditis; infection; metabolic disorder; obesity;  
KW reproductive disorder; infertility; neurological disorder;  
KW Parkinson's disease; Alzheimer's disease; stroke; epilepsy;  
KW cardiovascular disorder; myocardial infarction; hypertension;  
KW eye disorder; cell proliferative disease; cancer; atherosclerosis;  
KW hepatitis.  
XX  
OS Homo sapiens.  
XX  
PN WO2003072729-A2.  
XX  
PD 04-SEP-2003.  
XX  
PF 21-FEB-2003; 2003WO-0505478.  
XX  
PR 22-FEB-2002; 2002US-0359513P.  
PR 19-MAR-2002; 2002US-0365795P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Sprague WW, Kable AE, Lee S, Griffin JA, Becha SD, Lee EA;  
PI Hafalia AJA, Khare R, Emerling BM, Marquis JP, Ramkumar J;  
PI Elliott VS, Richardson TW, Baughn MR, Jin P, Chien D, Hawkins PR;  
PI Wilson AD, Chawla NK, Tran UK, Lee SY, Zebardjian Y, Jiang X;  
PI Jackson AA, Bhatia UG, Burrill JD, Blake JJ, Ho A, Zheng W;

XX WPI; 2003-731608/69.  
DR N-PSDB; ADD12630.  
XX  
PT New human enzymes and polynucleotides, useful for diagnosing, preventing  
PT or treating diseases or conditions associated with aberrant enzyme  
PT expression, e.g. cancer, hepatitis, AIDS, atherosclerosis, obesity,  
PT infection or stroke.  
XX  
PS Claim 1; SEQ ID NO 33; 426pp; English.  
XX  
CC The present sequence represents a human enzyme designated ENZM-33. Human  
CC enzymes of the present invention have cytostatic, antiarteriosclerotic,  
CC anti-HIV, antiallergic, cerebroprotective, antiparkinsonian,  
CC anticonvulsant, nontropic, neuroprotective, antiinflammatory,  
CC ophthalmological, antithyroid, antirheumatic, antibacterial, virucide,  
CC protozoacide, antiparasitic, fungicide, anorectic, cardiant, hypotensive,  
CC antinfertility and hepatotropic activities, and can be used in gene  
CC therapy. The human enzymes and polynucleotides encoding them can be used  
CC in diagnosing, preventing or treating diseases or conditions associated  
CC with the decreased expression or overexpression of the enzymes, such as  
CC autoimmune/inflammatory disorders (AIDS, allergies, atopic dermatitis,  
CC arthritis or thyroiditis), infections (e.g. bacterial, viral, parasitic,  
CC protozoal or fungal), metabolic disorders (e.g. obesity), reproductive  
CC disorders (e.g. infertility), neurological disorders (Parkinson's  
CC disease, Alzheimer's disease, stroke, epilepsy), cardiovascular disorders  
CC (e.g. myocardial infarction or hypertension), eye disorders, or cell  
CC proliferative diseases (e.g. cancer, atherosclerosis, hepatitis). They  
CC are also useful in assessing the effects of exogenous compounds on the  
CC expression of nucleic acid and amino acid sequences of the enzymes.  
XX  
SQ Sequence 310 AA;  
Query Match 28.9%; Score 442; DB 7; Length 310;  
Best Local Similarity 31.5%; Pred. No. 3e-37;  
Matches 94; Conservative 61; Mismatches 105; Indels 38; Gaps 4;  
QY 18 KYFPHGVRLPPFCRGKMEETANFVRPSDVIVTPKSGTSLLOEVVLYSQGADPDEI 77  
DB 12 KLKEVEGTLQPATVDNMSQIQSFEAKPDDLICTYKAGTWTWIEIVDMTEQNDVEKC 71  
QY 78 GLMNIDEOLPVLEY---POP-----GLDIIKELTSPRIKSHLPYRFLPS 119  
DB 72 QRALIQHRHPFIEWARPPQPSVULVRCFLSYFGVEKAKAMPSPRIKTHLSTQLLPP 131  
QY 120 DLHNGDSKVIYMARNPDLVVSYYQFHRSLRTMSYRGTFQFCRRFNMNDKLGYSWFHRV 179  
DB 132 SFWENNCKFLYVARNAKDCMVSYYHFQRMNHLDPDGTWEEYFETFFGVWGSWFDHV 191  
QY 180 QEFWEHRMDSNVFLKYEDMHRDLVTMVEQLARFLGVSCDKAOLEALTEH----- 229  
DB 192 KGWEMKDRHQILFLFYEDIKRDPKHEIRKVMQFMKKVDETVLDKIVQETSFKMKENP 251  
QY 230 -----CHQLVDOCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLYVYKQKVKCDLTF 280  
DB 252 MTRNSTVSKSILDQSISSF---MRKGTVDGNHFTVAQNERFDEIYRRKMEGTSINF 306  
RESULT 53  
ADA10983  
ID ADA10983 standard; protein; 296 AA.  
XX  
AC ADA10983;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE Human cDNA differentially expressed in colon cancer #68 product.  
XX  
KW differential expression; colon cancer; cancer; human.  
XX  
OS Homo sapiens.  
XX  
PN US2002160382-A1.

```

XX PD 31-OCT-2002.
XX PF
XX PF 11-OCT-2001; 2001US-00981353.
XX PR 11-OCT-2000; 2000US-0239841P.
XX PR (LASE/) LASEK A W.
XX PA (JONE/) JONES D A.
XX PI Lasek AW, Jones DA;
XX XX WPI; 2003-265756/26.
DR N-PSDB; ADA10982.
XX PT
XX PT New combination comprising cDNAs that are differentially expressed in
XX PT colon disorder, useful for diagnosing, treating, staging or monitoring
XX PT treatment for colon cancers.
XX PS Example 14; SEQ ID NO 101; 231pp; English.
XX CC The invention relates to a combination comprising cDNAs that are
XX CC differentially expressed in colon disorder. The methods and compositions
XX CC of the present invention are useful for diagnosing, treating, staging or
XX CC monitoring treatment for colon cancer. They are also useful in high
XX CC throughput methods for using cDNAs to detect differential expression of
XX CC nucleic acids in a sample, screening molecules or compounds to identify a
XX CC ligand which specifically binds a cDNA and using a protein to screen
XX CC molecules or compounds to identify at least one ligand which specifically
XX CC binds the protein. The present sequence represents the amino acid
XX CC sequence of a human cDNA differentially expressed in colon cancer
XX CC protein.
XX SQ Sequence 296 AA;
Query Match 28.8%; Score 440.5; DB 6; Length 296;
Best Local Similarity 33.9%; Pred. No. 4.1e-37;
Matches 94; Conservative 54; Mismatches 110; Indels 19; Gaps 3;
QY 23 HGVRLPPFCRGKMEBIANFPVRPSDVWIVTPKSGTSLLOEVVYLVSQADPDEIGLMNI 82
DB 16 HGYPMTCAFASWKEIQFHSRPDDIVATYPKSGTTWVSEIIDMLNDGDIKCKRGFI 75
QY 83 DEQLPVLEYPQGL-----DIKELTSPRLIKSHLPYRFLPSDLHNGSKVIYMARPKD 137
DB 76 TEKVPMEMLTLPGLRTSGTEQLKPNPSPIVKTHLFTDLLPKSFWENCKMIYLARNKD 135
QY 138 LVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVOEFWEHRMDSNVLFKAYE 197
DB 136 VSVSYTHFDLMNLPQFPQTWEYLEKFLTGKVAIGSWFTHVKWKKKEGHPILFYIE 195
QY 198 DMHRLDVTMVEQLARFLGVSQDKAQLAEALTEHCHLV---DQCCNAEALP----- 244
DB 196 DMKENPKBEIKIIRFLEKNLNDLRIIHTSFVEMKDNPLVNYTHLPTTVMDSKSP 255
QY 245 -VGRGRVGLWKDIPFVSMNEKFDLVYKQMGKCDLTF 280
DB 256 FMRKGTAGDKNYFTVAQNEKFDATYETEMSKTALQF 292
RESULT 54
AAW40498
ID AAW40498 standard; protein; 294 AA.
XX AC AAW40498;
XX AC
XX DT 14-JUL-1998 (first entry)
XX DE Human EST protein.
XX XX
XX KW Oestrogen sulphotransferase; EST; drug metabolism; xenobiotic; hormone;
XX KW neurotransmitter; sulphate conjugation; renal excretion; screening;
XX KW antigen; immunohistochemistry.

```

```

XX OS Homo sapiens.
XX PN US5744355-A.
XX PD 28-APR-1998.
XX PF
XX PF 09-MAY-1995; 95US-00437795.
XX PR 18-OCT-1994; 94US-00325562.
XX PA (MAYO-) MAYO FOUNDATION.
XX PI Wood TC, Aksoy IA, Weinshilboum RM;
XX DR WPI; 1998-271072/24.
XX DR N-PSDB; AAV11154.
XX PT Cells and vectors encoding human oestrogen sulpho-transferase - are
XX PT useful for the recombinant expression of the enzyme.
XX PS Claim 2; Fig 2; 30pp; English.
XX CC This sequence represents a novel human oestrogen sulphotransferase (EST).
XX CC This protein can be inserted into cell lines or vectors which are useful
XX CC for the recombinant expression of the enzyme. Sulphotransferases are
XX CC involved in the metabolism of drugs, xenobiotics, neurotransmitters and
XX CC hormones by sulphate conjugation. This is required for increase in water
XX CC solubility (leading to renal excretion) and in some cases for activation
XX CC of these compounds. The recombinant products are useful for the
XX CC determination of hormones and drugs which are metabolised by the enzyme
XX CC and screening of compounds structurally similar to oestrogen that can be
XX CC metabolised by the in vivo sulphotransferase and their affinity for the
XX CC enzyme. DNA encoding the enzyme can be used to quantify expression levels
XX CC in different tissues and detect mutations. The enzyme can also be used as
XX CC an antigen to raise antibodies for immunohistochemical analysis
XX SQ Sequence 294 AA;
Query Match 28.4%; Score 434; DB 2; Length 294;
Best Local Similarity 33.3%; Pred. No. 1.9e-36;
Matches 95; Conservative 56; Mismatches 116; Indels 18; Gaps 3;
QY 14 EPSEKYFEFHGVRLLPPFCRGKMEBIANFPVRPSDVWIVTPKSGTSLLOEVVYLVSQAD 73
DB 6 DYKFKFEVHGILMYKOFVKYWDNVEAFQARPDLDVIATYPKSGTTWVSEIVMIYKEGD 65
QY 74 PDEIGLMNIDQLPVLEYPQP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVI 129
DB 66 VEKCKEDVFNRIIPFLECRKENLNGVKQLDEMNSPRIVKTHLPPELLPASFWEKDCKII 125
QY 130 YMARNPDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVOEFWEHRMDS 189
DB 126 YLCRNAKDVAVSFYFFFLMWAGHPGSPFPEFVEKFMQGVPGYSWYKHKVSMWEKSKSP 185
QY 190 NVLFLKYEDMHRDLVTMVEQLARFLGVSQDKAQLAEALTEHCH-----QLVDQ 236
DB 186 RVLFVYEDLKEDIRKEVIKLIHFLERKPSBELVDRIIHTSFQEMKNPNSTNYTTLDPDE 245
QY 237 CCNAEALP -VGRGRVGLWKDIPFVSMNEKFDLVYKQMGKCDLTF 280
DB 246 IMNQKLSPPMRKGTGDWKNHFTVALNEKFDKHYEQQMKESTLKF 290
RESULT 55
AAW44247
ID AAW44247 standard; protein; 294 AA.
XX AC AAW44247;
XX AC
XX DT 15-MAY-1998 (first entry)
XX DE Human oestrogen sulphotransferase.

```

```
XX Human; oestrogen sulphotransferase; estrogen sulphotransferase; liver EST;
KW biotransformation.
XX
XX Homo sapiens.
XX
XX US5714594-A.
XX
XX 03-FEB-1998.
XX
XX 18-OCT-1994; 94US-00325562.
XX
XX 18-OCT-1994; 94US-00325562.
XX
XX (MAYO-) MAYO FOUNDATION.
XX
XX Wood TC, Weinshilboum RM, Aksoy IA;
XX
XX WPI; 1998-144284/13.
XX
XX N-PSDB; AAV12466.
XX
XX DNA encoding human oestrogen sulphotransferase - useful for research into
XX oestrogen biotransformation.
XX
XX Claim 1; Col 21-24; 25pp; English.
XX
XX The present sequence represents a human oestrogen sulphotransferase (EST)
XX protein. The isolation and expression of cDNA which encodes human liver
XX EST enables the biotransformation of oestrogens to be studied
XX
XX Sequence 294 AA;
XX
XX Query Match 28.4%; Score 434; DB 2; Length 294;
XX Best Local Similarity 33.3%; Pred. No. 1.9e-36;
XX Matches 95; Conservative 56; Mismatches 116; Indels 18; Gaps 3;
XX
XX 14 EFESKYFEFHGVRLPPFCRCGRKMEETANFPVRPSDVWIVTPKSGTSLLOEVVYLVSQAD 73
XX 6 DYKFEFEEVHGILMYKDFVKYWDNVEAFQARDDLVVATPKSGTTWVSEIVYMYIKREGD 65
XX
XX 74 PDEIGLNMNIDQLPVLVEYPOP-GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVI 129
XX 66 VEKCKEDVFNRIPELECKENLNGVKQLDEMNSPRIVKTHLPPELLPASFWKDCCKII 125
XX
XX 130 YMARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDS 189
XX 126 YLCRNAKDVAVSFYFFFLMVAGHPNPGSPFPEFVEKFMQGPYGSWYKHKVSWWEKKGKSP 185
XX
XX 190 NVLEFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLALETCH-QLVDQ 236
XX 186 RVLFYFEDLKBEDIRKEVIKLHFLERKPSSELVDRIIHTTSFQEMKNPNSTNYTTLDPDE 245
XX
XX 237 CCNAEALF-VGRGRVGLMKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
XX 246 IMNQKLSFPMKGIITGDKNHETVALNEKFDXHEQQMKESTLKF 290
XX
XX RESULT 56
XX ID ADD48896
XX
XX ADD48896 standard; protein; 294 AA.
XX
XX AC ADD48896;
XX
XX 02-DEC-2004 (revised)
XX 29-JAN-2004 (first entry)
XX
XX Human Protein P49888, SEQ ID NO 14607.
XX
XX Human; pain; neuronal tissue; gene therapy;
XX spinal segmental nerve injury; chronic constriction injury; CCI;
XX spared nerve injury; SNI; Chung.
XX
XX Homo sapiens.
XX
```

```
OS Unidentified.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
XX
XX 01-NOV-2001; 2001US-0346382P.
XX
XX 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO ) GEN HOSPITAL CORP.
XX
XX (FARB ) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
XX
XX GENBANK; P49888.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Example 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a human protein (described in Table 3
XX of the specification) which is differentially expressed during pain.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 294 AA;
XX
XX Query Match 28.4%; Score 434; DB 7; Length 294;
XX Best Local Similarity 33.3%; Pred. No. 1.9e-36;
XX Matches 95; Conservative 56; Mismatches 116; Indels 18; Gaps 3;
XX
XX 14 EFESKYFEFHGVRLPPFCRCGRKMEETANFPVRPSDVWIVTPKSGTSLLOEVVYLVSQAD 73
XX 6 DYKFEFEEVHGILMYKDFVKYWDNVEAFQARDDLVVATPKSGTTWVSEIVYMYIKREGD 65
XX
XX 74 PDEIGLNMNIDQLPVLVEYPOP-GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVI 129
XX 66 VEKCKEDVFNRIPELECKENLNGVKQLDEMNSPRIVKTHLPPELLPASFWKDCCKII 125
XX
XX 130 YMARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDS 189
XX 126 YLCRNAKDVAVSFYFFFLMVAGHPNPGSPFPEFVEKFMQGPYGSWYKHKVSWWEKKGKSP 185
XX
XX 190 NVLEFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLALETCH-QLVDQ 236
XX 186 RVLFYFEDLKBEDIRKEVIKLHFLERKPSSELVDRIIHTTSFQEMKNPNSTNYTTLDPDE 245
XX
XX 237 CCNAEALF-VGRGRVGLMKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
XX 246 IMNQKLSFPMKGIITGDKNHETVALNEKFDXHEQQMKESTLKF 290
XX
```

Db	186	RVLFLFYEDLKDIREKVIKLIHFLERKPSBELVDRIIHHTSFQBMKNPNSTNYTTLDPDE	245	
Qy	237	CCNAEALP-VGRGRVLGWLKIDFTVSMNEKFDLIVYKQKMGKCDLTF	280	
Db	246	IMNQKLSPFMRKGTGDWKNHFTVALNEKFKDKHVEQOMKESLKF	290	
RESULT 57				
AD117049				
ID	AD117049	standard; protein; 307 AA.		
XX	AC	AD117049;		
XX	DT	15-APR-2004 (first entry)		
XX	DE	Chicken NOVX protein homologue SeqID 585.		
XX	KW	chicken; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;		
XX	KW	inflammation; autoimmune disorder; allergy; blood disorder;		
XX	KW	acquired immunodeficiency syndrome; AIDS; obesity; asthma;		
XX	KW	immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;		
XX	KW	Alzheimer's disease; infection; s.		
OS	Gallus gallus.			
XX	PN	WO200268649-A2.		
XX	PD	06-SEP-2002.		
XX	PF	31-JAN-2002; 2002WO-US002785.		
XX	PR	31-JAN-2001; 2001US-0265339P.		
PR	31-JAN-2001; 2001US-0265412P.			
PR	31-JAN-2001; 2001US-0265514P.			
PR	31-JAN-2001; 2001US-0265517P.			
PR	02-FEB-2001; 2001US-0266406P.			
PR	05-FEB-2001; 2001US-0266767P.			
PR	07-FEB-2001; 2001US-0266975P.			
PR	07-FEB-2001; 2001US-0267057P.			
PR	08-FEB-2001; 2001US-0267459P.			
PR	09-FEB-2001; 2001US-0267823P.			
PR	15-FEB-2001; 2001US-0268974P.			
PR	26-FEB-2001; 2001US-0271664P.			
PR	27-FEB-2001; 2001US-0271839P.			
PR	27-FEB-2001; 2001US-0271855P.			
PR	02-MAR-2001; 2001US-0272788P.			
PR	02-MAR-2001; 2001US-0273046P.			
PR	14-MAR-2001; 2001US-0275925P.			
PR	14-MAR-2001; 2001US-0275947P.			
PR	14-MAR-2001; 2001US-0275950P.			
PR	14-MAR-2001; 2001US-0275989P.			
PR	15-MAR-2001; 2001US-0276448P.			
PR	15-MAR-2001; 2001US-0276450P.			
PR	16-MAR-2001; 2001US-0276397P.			
PR	16-MAR-2001; 2001US-0276768P.			
PR	20-MAR-2001; 2001US-0278652P.			
PR	26-MAR-2001; 2001US-0278775P.			
PR	26-MAR-2001; 2001US-0278778P.			
PR	29-MAR-2001; 2001US-0279882P.			
PR	30-MAR-2001; 2001US-0279884P.			
PR	30-MAR-2001; 2001US-0280147P.			
PR	11-APR-2001; 2001US-0282992P.			
PR	11-APR-2001; 2001US-0283083P.			
PR	20-APR-2001; 2001US-0285133P.			
PR	23-APR-2001; 2001US-0285749P.			
PR	03-MAY-2001; 2001US-0288327P.			
PR	03-MAY-2001; 2001US-0288504P.			
PR	29-MAY-2001; 2001US-0294047P.			
PR	30-MAY-2001; 2001US-0294473P.			
PR	08-JUN-2001; 2001US-0296964P.			
PR	18-JUN-2001; 2001US-0298959P.			
PR	19-JUN-2001; 2001US-0299324P.			
PR	13-AUG-2001; 2001US-0312020P.			
PR	16-AUG-2001; 2001US-0312889P.			
PR	16-AUG-2001; 2001US-0312908P.			
PR	21-AUG-2001; 2001US-0313390P.			
PR	28-AUG-2001; 2001US-0315470P.			
PR	31-AUG-2001; 2001US-0316447P.			
PR	07-SEP-2001; 2001US-0318115P.			
PR	07-SEP-2001; 2001US-0318118P.			
PR	12-SEP-2001; 2001US-0318740P.			
PR	19-SEP-2001; 2001US-0323379P.			
PR	18-OCT-2001; 2001US-0330245P.			
PR	18-OCT-2001; 2001US-0330308P.			
PR	14-NOV-2001; 2001US-0332701P.			
XX	PA	(CURA-) CURAGEN CORP.		
XX	PI	Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;		
PI	Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;			
PI	Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CEA;			
PI	Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;			
XX	WPI; 2002-706998/76.			
XX	New NOVX polypeptides and nucleic acids, useful for preventing or			
PT	treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,			
PT	atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or			
PT	pharmacogenomics.			
XX	Disclosure; SEQ ID NO 585; 1498pp; English.			
XX	This invention relates to a novel nucleic acids, and encoded polypeptides			
CC	thereof, which have properties related to the stimulation of biochemical			
CC	or physiological responses in a cell, tissue, organ or organism.			
CC	Specifically, it refers to the use of biologically active fragments for			
CC	diagnostic and prognostic assays and furthermore in the treatment of			
CC	diverse pathological conditions. The present invention describes novel			
CC	human and murine NOVX proteins, as well as methods to modulate their			
CC	expression using antisense oligos, ribozymes and peptide nucleic acids			
CC	The NOVX polypeptides, polynucleotides and antibodies are useful in			
CC	treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,			
CC	atherosclerosis, cancer and diabetes. Furthermore, they may be used in			
CC	treating or preventing diseases such as inflammation, autoimmune			
CC	disorders, allergies, blood disorders, acquired immunodeficiency syndrome			
CC	(AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,			
CC	arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy			
CC	and epilepsy. Accordingly, these molecules have many activities including			
CC	cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,			
CC	haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,			
CC	antiasthmatic, nephrotropic, antiarthritic, virucide, hepatotropic,			
CC	neuroprotective, neurotropic, antibacterial, virucide, antiparasitic,			
CC	relaxant and anticonvulsant. In addition, they are useful in screening			
CC	assays to identify small molecules that modulate or inhibit, for example,			
CC	neurogenesis, wound healing and angiogenesis. The nucleic acids are also			
CC	used as in chromosome mapping, tissue typing, preventive medicine and			
CC	pharmacogenomics. This polypeptide is a homologue of a human NOVX protein			
CC	of the invention.			
XX	Sequence 307 AA;			
SQ	Query Match 28.0%; Score 428; DB 5; Length 307; Best Local Similarity 32.6%; Pred. No. 8.8e-36; Matches 87; Conservative 61; Mismatches 97; Indels 22; Gaps 2;			
Qy	36	BEIANFPVRPSDWMIVITPKSGTSLQLQEVVLVVSQGDPPDEIGLMNIDEQIPVLEYPOP- 94		
Db	37	DQWNFKARPDLLVATYKAGTTTQBEIVDMIQQNGDIEKCRRASTYKRPFLFLEWYIPD 96		
Qy	95	-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDKLVSYQFHR 147		
Db	97	SSPLGYSLKLAEMPSPRTNKTHTLPVLQVPSPFWEQNCIIYVARNAKDLVSYFHR 156		
Qy	148	SLRTMSYRGTFQFCRRPMDNKLGVGSWFHFQVWEHRMDSNVLFLEYDMHRLVTMV 207		
Db	157	MNKVLPDPGTTEETFEKFMNGEVLWGSWYDHVKGWKAQKDKHRIILYLFYEDMKENPKREI 216		



XX PS Claim 25; Fig 2; 103pp; English.

XX CC The invention relates to a novel human sulfotransferase polypeptide. The

XX CC sulfotransferase has cytosolic activity. The polynucleotide of the

XX CC invention may have a use in gene therapy. The polypeptide and

XX CC polynucleotide of the invention are useful in preventing, ameliorating,

XX CC or correcting diseases including cancer and chronic obstructive pulmonary

XX CC disease (COPD). The sequence represents a human sulfotransferase of the

XX CC invention

SQ Sequence 304 AA;

Query Match 27.6%; Score 422; DB 5; Length 304;

Best Local Similarity 32.3%; Pred. No. 3.7e-35;

Matches 97; Conservative 59; Mismatches 124; Indels 20; Gaps 5;

QY 1 MAESAETPSTGEPE-SKYFEFHGVRLPPFCRGKMEETANPPVRPSDVWIVTPKSGTS 59

DB 1 MAKIEKNAPTWEKKPELFINMEVDGPTLILSKEWKVCNFOAKPDDLILATYPKSGTT 60

QY 60 LQEVVYLVSGADPDEIGLMNIDQLPVL-YP---QPGLDIIKELTSPRLIKSHLPY 114

DB 61 WMHEILMILNDGDVEKCKRAQTLDRHAFLELKPFPKEKPDLEFVLEMSSPOLIKTHLPS 120

QY 115 RFLPSDLHNGSKVIYMARPNKDLVSVYQPHRSRLTWSYRGTFQFCRRFMDKLGYS 174

DB 121 HLIPPSIWKENCKIYVARNPKDCLVSYHFRHMASFMPDPQNLEEFYEKFMKGVGGS 180

QY 175 WFEHQVEFWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCDAQLEALTEHQVLV 234

DB 181 WFDVKGHWAAKDMRILYLFYEDIKDKPKREIKILFLEKDISEILNKIIYHTSPDV 240

QY 235 ---DQCCNAEALPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280

DB 241 MKQNPMTWNTLTPTSIMDHSISPPFMRKGMPGDKWKNYFTVAQNEEDKDYQKMGAGTLTF 300

RESULT 60

ID ADD48894

XX ADD48894 standard; protein; 295 AA.

XX AC ADD48894;

XX DT 02-DEC-2004 (revised)

XX DT 29-JAN-2004 (first entry)

XX DE Rat Protein P52844, SEQ ID NO 14605.

XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

XX KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX OS Unidentified.

XX PN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GEHO ) GEN HOSPITAL CORP.

XX PA (FARB ) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX DR GENBANK; P52844.

PT New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX Example 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat

XX CC or human polynucleotides or a polynucleotide which represents a fragment,

XX CC derivative or allelic variation of the nucleic acid sequence. Also

XX CC claimed are a vector comprising the novel polynucleotide, a host cell

XX CC comprising the vector, a method for identifying a nucleotide sequence

XX CC which is differentially regulated in an animal subjected to pain and a

XX CC kit to perform the method, an array, a method for identifying an agent

XX CC that increases or decreases the expression of the polynucleotide sequence

XX CC that is differentially expressed in neuronal tissue of a first animal

XX CC subjected to pain, a method for identifying a compound which regulates

XX CC the expression of a polynucleotide sequence which is differentially

XX CC expressed in an animal subjected to pain, a method for identifying a

XX CC compound that regulates the activity of one or more of the

XX CC polynucleotides, a method for producing a pharmaceutical composition, a

XX CC method for identifying a compound or small molecule that regulates the

XX CC activity in an animal of one or more of the polypeptides given in the

XX CC specification, a method for identifying a compound useful in treating

XX CC pain and a pharmaceutical composition comprising the one or more

XX CC polypeptides or their antibodies. The polynucleotide or the compound that

XX CC modulates its activity is useful for preparing a medicament for treating

XX CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

XX CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene

XX CC therapy). The sequence presented is a rat protein (described in Table 3

XX CC of the specification) which is differentially expressed during pain.

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic form directly from WIPO at

XX CC ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 295 AA;

Query Match 27.5%; Score 421; DB 7; Length 295;

Best Local Similarity 31.9%; Pred. No. 4.5e-35;

Matches 91; Conservative 60; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYFEFHGVRLPPFCRGKMEETANPPVRPSDVWIVTPKSGTSLQEVVYLVSGAD 73

DB 7 EYEVFVGFHGLMDKRFYKWEIDETFLARPDDLIVTPKSGSTWISIVDMYIKRGD 66

QY 74 PDEIGLMNIDQLPVLVYQPP-GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVI 129

DB 67 VEKCKEDALFNRIIDLECRNEDLINGIKQLKEKSPRIIVKTHLPAKLPAFWEKKNCKII 126

QY 130 YMARNPKDLVSVYQPHRSRLTWSYRGTFQFCRRFMDKLGYSWFEHQVEFWEHRMDS 189

DB 127 YLCRNADVVVYVYFFFLIMKSYNPKSFSEFVEKFMGQVPYGSWYDHVKSWEKSKNS 186

QY 190 NVLFLKYEDMHRDLVTMVEQLARFLGVSCDAQLEALTEHC---HQLVDQCCNAEALP-- 244

DB 187 RVLFMFYEDMKEDIRREVVKLIEFLERDPSAELVDRIIQTSTFQEMKKNPCTNYSMLPET 246

QY 245 -----VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280

DB 247 MIDLVSPFMRKGIWGDWRNHFPEALRFRFBEHYQRHMKOCVPKF 291

RESULT 61

ABB81793

ID ABB81793 standard; protein; 304 AA.

XX ABB81793;

XX AC ABB81793;

XX DT 24-SEP-2002 (first entry)

XX DE Human sulfotransferase #3.

XX KW Human; sulfotransferase; cytostatic; gene therapy; cancer; COPD;

XX KW Chronic obstructive pulmonary disease.











CC for constructing recombinant vectors and host cells, expressing antigenic  
CC portions of the proteins, in genotyping and mapping experiments, for  
CC designing ribozymes, for monitoring the effectiveness of modulating  
CC compounds on the expression or activity of the protein, in diagnostic  
CC assays, antisense constructs to control protein expression and in gene  
CC therapy. The present sequence is ST drug-metabolising protein 3 encoded  
CC by cDNA  
XX  
SQ Sequence 304 AA;

Query Match 26.6%; Score 407; DB 4; Length 304;  
Best Local Similarity 31.3%; Pred. No. 1.4e-33;  
Matches 94; Conservative 61; Mismatches 125; Indels 20; Gaps 5;  
QY 1 MAESAETPTSGEPE-SKYFEFHGVRLLPPFCRGKMEETIANFPVRPSDWMIVTPKSGTS 59  
DB 1 MAKIEKNAPTMEKKPELFNIMEVDGVPITLLSKWEKWCNFOAKPDOLLATYPKSGTT 60  
QY 60 LLOEVVYLVSQADPEIGLNMIDQLPVLE--YP---QPGLDIIKELTSPRIKSHLPY 114  
DB 61 WHEILDMILNDGDVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQIILHLP 120  
QY 115 RFLPSDLHNGDSKVITYMARNPXDLVVSYYQFHRSLRTWSYRGTFQFCRRFMNDKLGYS 174  
DB 121 HLIPPSINKCKIYVVARNPXDLVSYHFRMASFPDPQNLSEFEYKFKMSGVVGS 180  
QY 175 WFEHQVEFWEHRMDSNVLFKLYEDMHRDLVTMVEQLARFLGVS CDKAQLEALTEHCHOLV 234  
DB 181 WFDVKGWAAKADTHRIYLFYEDIKNPKHEIKHVLFEKLTSGDVINKIVHTSPDV 240  
QY 235 ---DQCCNAELP-----VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280  
DB 241 MKDNPMANHTAVPAHIFNHSISKFRKMGPMGDWKNHFTVAMNENFDKHYEKKMGSTLNF 300

RESULT 67  
AB866006  
ID AB866006 standard; protein; 316 AA.  
XX  
AC AB866006;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 24810.  
XX  
DE Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
KW Drosophila melanogaster.  
OS  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
DR N-PSDB; ABL10109.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.  
XX  
PS Disclosure; SEQ ID NO 24810; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 316 AA;

Query Match 26.5%; Score 405; DB 4; Length 316;  
Best Local Similarity 34.4%; Pred. No. 2.4e-33;  
Matches 94; Conservative 52; Mismatches 99; Indels 28; Gaps 8;  
QY 36 BEIANFPVRPSDWMIVTPKSGTSLLOQVYLVLSQ-----GADPD-----EI-GLMN 81  
DB 45 ESIRSLPVYQDDVMVSYPRGTGWAEQWVLLGHQLDVYAAEQDLRLRSPLIELSALFS 104  
QY 82 IDEQLPVLEYQPGGLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVITYMARNPXDLVVS 141  
DB 105 IDHETVAQKFGNTVDLVNLPFRFARSHLPWLLPQFETVKPRIVVYTARNPKDLCVS 164  
QY 142 YYQFHRSLRTWSYRGTFQFCRRFMNDKLGYSWFEHQVEFWEHRMDSNVLFKLYEDMHR 201  
DB 165 YYHYFKLLHGMN--GDPEQFVDLFLEGHTPMGYSYWRHVLFPWKRSDQDNVLFKYE 222  
QY 202 DLVTMVEQLARFLGVS--CDKAQLEALTEHC--HQLVDOCCNAE-ALP-----VGRGR 249  
DB 223 DLPSVVRRCARPLGVQSLLDVSTQKCDHITFDKMRANKAVNLEKLPSESSKFIRNGK 282  
QY 250 VGLMKDIFTVSMNEKFDLVYKQMGKCDLTFDF 282  
DB 283 IGDWRNHNHNMESRSEFDEWTERHMRGSGLNFDY 315

RESULT 68  
AAE12208  
ID AAE12208 standard; protein; 304 AA.  
XX  
AC AAE12208;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Human ST drug-metabolising protein 1 encoded by DNA transcript 1.  
KW Human; sulfoltransferase; drug-metabolising enzyme mediated disorder; ST;  
KW DME; toxicity study; gene therapy.  
OS Homo sapiens.  
XX  
FH Location/Qualifiers  
FT Modified-site 173..175  
FT /note= "Protein kinase C phosphorylation site"  
FT Modified-site 180..183  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 236..239  
FT /note= "Casein kinase II phosphorylation site"  
FT Modified-site 248..251  
FT /note= "Asn is N-glycosylated"  
FT Modified-site 258..261  
FT /note= "Asn is N-glycosylated"  
FT Modified-site 295..300  
FT /note= "N-myristoylation site"  
XX  
PN WO200172977-A2.  
XX  
PD 04-OCT-2001.  
XX  
PF 27-MAR-2001; 2001WO-US009659.  
XX  
PR 27-MAR-2000; 2000US-0192408P.

PR 20-JUN-2000; 2000US-0212725P.  
PR 03-JUL-2000; 2000US-00609816.  
XX (PEKE ) PE CORP NY.  
PA  
XX Woodage T, Wei M, Kodira C, Di Francesco V, Beasley EM;  
XX WPI; 2001-616484/71.  
XX N-P5DB; RAD19776.  
DR  
DR  
XX  
PT New polypeptides useful in labeling reagents with tagged sulfur and  
PT identifying modulators of the protein comprises proteins related to  
PT sulfotransferase drug-metabolizing enzyme subfamily.  
XX  
XX Claim 1; Fig 2; 83pp; English.  
XX  
CC The invention relates to sulfotransferase (ST) drug-metabolizing enzyme  
CC (DME) peptides, proteins and nucleic acid molecules encoding them.  
CC Sequences of the invention are useful in labelling reagents with tagged  
CC sulphur, for the development of human therapeutics, as query sequences to  
CC perform a search against sequence databases, for identifying other family  
CC members or related sequences, to raise antibodies or to elicit another  
CC immune response, as a reagent in assays designed to quantitatively  
CC determine levels of the protein in biological fluids, to metabolise  
CC compounds for toxicity studies, and as markers for tissues in which the  
CC corresponding protein is preferentially expressed. Polypeptides of the  
CC invention is also useful in diagnosing a disease, predisposition to a  
CC disease mediated by the peptide, in pharmacogenomic analysis and treating  
CC a disorder characterised by absence of inappropriate or unwanted  
CC expression of the protein. Polynucleotides of the invention are useful  
CC for constructing recombinant vectors and host cells, expressing antigenic  
CC portions of the proteins, in genotyping and mapping experiments, for  
CC designing ribozymes, for monitoring the effectiveness of modulating  
CC compounds on the expression or activity of the protein, in diagnostic  
CC assays, antisense constructs to control protein expression and in gene  
CC therapy. The present sequence is ST drug-metabolizing protein 1 encoded  
CC by DNA transcript 1  
XX  
SQ Sequence 304 AA;  
Query Match 26.2%; Score 401; DB 4; Length 304;  
Best Local Similarity 31.0%; Pred. No. 5.9e-33;  
Matches 93; Conservative 62; Mismatches 125; Indels 20; Gaps 5;  
QY 1 MAESAETPSTGEFE-SKYFEFHGVLPPFCRGKMEETANFPVPSDVWIVTPKSGTS 59  
DB 1 MAKIEKNAPTMEKKPELFNIMEVDGVPTLLISKWEWKEVCNFOAKPDDLILATYPSGTT 60  
QY 60 LLQEVVYLVSQGADPEIGLNMIDQLPVL--YP---QPGLDIIKELTSPRLIKSHLPY 114  
DB 61 WMHEILDMINDGDVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQLIKTHLPS 120  
QY 115 RFLPSDLHNGDSKVITYMARNPKDLVSVYQFHRSLRTWTSYRGTFOEFCRRFMNDKLGYS 174  
DB 121 HLIPPSIWKENCKIVYVARNPKDCLVSYVHFHRMASFMPDPQNLBEFYEKFGSKGVGGS 180  
QY 175 WFEHQVEWHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTECHQLV 234  
DB 181 WFDVHKWAAKMDHRIILYFLFYFKKPKKBEIHKVLEFLEKTSWGDVINKIVHRTSFDV 240  
QY 235 ---DOCCNAEALP-----VGRGRVGLWKDIFTVMNEKFDLVYKQWKGKCDLTF 280  
DB 241 MKDNPMANHTAVPAHFNHISKFMKRGMPGDKKHNFTVALNENFDKHYEKWAGSTLNF 300  
RESULT 69  
AD117046  
ID AD117046 standard; protein; 304 AA.  
XX  
AC AD117046;  
XX  
DT 15-APR-2004 (first entry)  
XX

DE Human NOVX protein homologue SeqID 582.  
XX  
KW human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;  
KW inflammation; autoimmune disorder; allergy; blood disorder;  
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;  
KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;  
KW Alzheimer's disease; infection; str.  
XX  
OS Homo sapiens.  
XX  
PN WO200268649-A2.  
XX  
PD 06-SEP-2002.  
XX  
PF 31-JAN-2002; 2002WO-US002785.  
XX  
PR 31-JAN-2001; 2001US-0265395P.  
PR 31-JAN-2001; 2001US-0265412P.  
PR 31-JAN-2001; 2001US-0265514P.  
PR 31-JAN-2001; 2001US-0265517P.  
PR 02-FEB-2001; 2001US-0266406P.  
PR 05-FEB-2001; 2001US-0266767P.  
PR 07-FEB-2001; 2001US-0266975P.  
PR 08-FEB-2001; 2001US-0267057P.  
PR 08-FEB-2001; 2001US-0267459P.  
PR 09-FEB-2001; 2001US-0267823P.  
PR 15-FEB-2001; 2001US-0268974P.  
PR 26-FEB-2001; 2001US-0271664P.  
PR 27-FEB-2001; 2001US-0271839P.  
PR 27-FEB-2001; 2001US-0271855P.  
PR 02-MAR-2001; 2001US-0272788P.  
PR 02-MAR-2001; 2001US-0273046P.  
PR 14-MAR-2001; 2001US-0275925P.  
PR 14-MAR-2001; 2001US-0275947P.  
PR 14-MAR-2001; 2001US-0275950P.  
PR 14-MAR-2001; 2001US-0275989P.  
PR 15-MAR-2001; 2001US-0276448P.  
PR 15-MAR-2001; 2001US-0276450P.  
PR 16-MAR-2001; 2001US-0276397P.  
PR 16-MAR-2001; 2001US-0276768P.  
PR 20-MAR-2001; 2001US-0278652P.  
PR 26-MAR-2001; 2001US-0278775P.  
PR 26-MAR-2001; 2001US-0278778P.  
PR 29-MAR-2001; 2001US-0279882P.  
PR 29-MAR-2001; 2001US-0279884P.  
PR 30-MAR-2001; 2001US-0280147P.  
PR 11-APR-2001; 2001US-0282992P.  
PR 11-APR-2001; 2001US-0283083P.  
PR 20-APR-2001; 2001US-0285133P.  
PR 23-APR-2001; 2001US-0285749P.  
PR 03-MAY-2001; 2001US-0288327P.  
PR 03-MAY-2001; 2001US-0288504P.  
PR 29-MAY-2001; 2001US-0294047P.  
PR 30-MAY-2001; 2001US-0294473P.  
PR 08-JUN-2001; 2001US-0296964P.  
PR 18-JUN-2001; 2001US-0298959P.  
PR 19-JUN-2001; 2001US-0299324P.  
PR 13-AUG-2001; 2001US-0312020P.  
PR 16-AUG-2001; 2001US-0312889P.  
PR 21-AUG-2001; 2001US-0312908P.  
PR 21-AUG-2001; 2001US-0313390P.  
PR 28-AUG-2001; 2001US-0315470P.  
PR 31-AUG-2001; 2001US-0316447P.  
PR 07-SEP-2001; 2001US-0318115P.  
PR 07-SEP-2001; 2001US-0318118P.  
PR 12-SEP-2001; 2001US-0318740P.  
PR 19-SEP-2001; 2001US-0323379P.  
PR 18-OCT-2001; 2001US-0330245P.  
PR 18-OCT-2001; 2001US-0330308P.  
PR 14-NOV-2001; 2001US-0332701P.  
XX  
XX (CURA-) CURAGEN CORP.

PI Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;  
PI Li L, Gangolli EA, Padigara M, Anderson DM, Rastelli L, Miller CE;  
PI Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CE;  
PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;  
XX WPI; 2002-706998/76.  
XX  
XX New NOVX polypeptides and nucleic acids, useful for preventing or  
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,  
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX  
PS Disclosure; SEQ ID NO 582; 1498pp; English.  
XX  
CC This invention relates to a novel nucleic acids, and encoded polypeptides  
CC thereof, which have properties related to the stimulation of biochemical  
CC or physiological responses in a cell, tissue, organ or organism.  
CC Specifically, it refers to the use of biologically active fragments for  
CC diagnostic and prognostic assays and furthermore in the treatment of  
CC diverse pathological conditions. The present invention describes novel  
CC human and murine NOVX proteins, as well as methods to modulate their  
CC expression using antisense oligos, ribozymes and peptide nucleic acids.  
CC The NOVX polypeptides, polynucleotides and antibodies are useful in  
CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,  
CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in  
CC treating or preventing diseases such as inflammation, autoimmune  
CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome  
CC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,  
CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy  
CC and epilepsy. Accordingly, these molecules have many activities including  
CC cyostatic, cardiant, antiinflammatory, immunosuppressive, anti allergic,  
CC haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,  
CC antiasthmatic, nephrotropic, antiarthritic, hepatotropic,  
CC neuroprotective, nootropic, antibacterial, viricide, antiparasitic,  
CC relaxant and anticonvulsant. In addition, they are useful in screening  
CC assays to identify small molecules that modulate or inhibit, for example,  
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also  
CC used as in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein  
CC of the invention.  
XX  
SQ Sequence 304 AA;  
  
Query Match 26.2%; Score 401; DB 5; Length 304;  
Best Local Similarity 31.0%; Pred. No. 5.9e-33;  
Matches 93; Conservative 62; Mismatches 125; Indels 20; Gaps 5;  
  
Qy 1 MAESAETPSTPGEFE-SKYFEFHGVRVLPFCRCGRKWEIANPVPSPDVWIVYPKSGTS 59  
Db 1 MAKIEKNAPTWEKKPELFNIMEVDGVPFTLLSKWEKVCNFKQAKPDDLLIATYPSGTT 60  
Qy 60 LQEVYVLVSQADPDEIGLNMIDQLPVLE--YP---QPGDIDIIKELTSPRLIKSHLPY 114  
Db 61 WHEILDMLNDGDVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEWSSPOLIKTHLPS 120  
Qy 115 RFLPSDLHNGSKVITYMARNKDLVSVYQFHSRLTWSYRGTFQFCRRFMNDKLGVGS 174  
Db 121 HLIPPSIWKENCKIYVAVARNKDLVSVYHFRMASFMPDPQNLBEFYEFKFWGSGVGS 180  
Qy 175 WFEHVQEFWEHMRDMSNVLFLKVEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHQVLV 234  
Db 191 WFDHVKGWAAKDMHRILYLFYEDIKKPKKHEIHKVLEFKTWGSDVINIKVHHSTFDV 240  
Qy 235 ---DOCCNAEALP-----VGRGRVGLWKDIFTVSMNEKFDLVYKQWKGKCDLTF 280  
Db 241 MKDNPMANHTAVPAHIFNHSISKFKRKGMPGDKKHNFTVALNENFDKHEKMGAGSTLNF 300  
  
RESULT 70  
ABU12099  
ID ABU12099 standard; protein; 305 AA.  
XX  
AC ABU12099;

XX 17-FEB-2003 (first entry)  
XX Novel human sulfotransferase-like protein.  
XX  
XX Gamma-aminobutyric acid receptor-like protein; depression; stroke;  
KW GABA receptor-like protein; Parkinson's disease; Huntington's disease;  
KW Tourette's syndrome; amyotrophic lateral sclerosis; head trauma;  
KW Alzheimer's disease; alcoholism; vigilance; anxiety; muscle tension;  
KW epileptogenic activity; memory; cardiomyopathy; cancer; angiogenesis;  
KW arrhythmogenic right ventricular dysplasia; renal disease; diabetes;  
KW Epidermal growth factor like protein; leukaemia; lupus; anaemia; ulcer;  
KW haematopoietic stem and progenitor cell like protein; cirrhosis;  
KW sulfotransferase-like protein; cholangitis; hepatitis; hyperthyroidism;  
KW developmental disorder; Syntaxin-like protein; myxoid liposarcoma;  
KW asthma; Lambert-Eaton myasthenic syndrome; acute myeloidleukaemia;  
KW transgenic animal.  
XX  
XX Homo sapiens.  
XX  
XX US2002123612-A1.  
XX  
XX 05-SEP-2002.  
XX  
XX 03-JUL-2001; 2001US-00898570.  
XX  
XX 19-APR-2000; 2000US-0198293P.  
XX 20-APR-2000; 2000US-0198645P.  
XX 25-APR-2000; 2000US-0199476P.  
XX 26-APR-2000; 2000US-0199880P.  
XX 26-APR-2000; 2000US-0200024P.  
XX 26-APR-2000; 2000US-0200025P.  
XX 09-JUN-2000; 2000US-0210809P.  
XX 03-JUL-2000; 2000US-0215855P.  
XX 17-JUL-2000; 2000US-0218591P.  
XX 11-AUG-2000; 2000US-0224610P.  
XX 27-FEB-2001; 2001US-0271814P.  
XX  
XX (GERL/) GERLACH V L.  
XX (ELLE/) ELLERMAN K.  
XX (MACD/) MACDOUGALL J R.  
XX (SMIT/) SMITHSON G.  
XX  
XX Gerlach VL, Ellerman K, Macdougall JR, Smithson G;  
PI WPI; 2003-066815/06.  
XX N-PSDB; ABX56478.  
XX  
XX Novel polypeptides and nucleic acids which are members of epidermal  
PT growth factor, complement receptor families for diagnosing and treating  
PT psychiatric conditions, depression, stroke, Alzheimer's and Parkinson's  
PT disease.  
XX  
XX Claim 1; Page 35; 91pp; English.  
XX  
XX The invention describes an isolated POLYX (POLY1-17) polypeptide and its  
CC variant. POLYX polypeptides (especially POLY5, POLY6 and POLY7), the  
CC polynucleotides encoding them (I) and an anti-POLYX-antibody (III) are  
CC useful for treating or preventing a pathology associated with POLYX  
CC polypeptide in humans and for treating a syndrome associated with human  
CC disease. POLYX polypeptide is also useful for identifying an agent that  
CC binds to POLYX and a cell expressing POLYX is useful for identifying a  
CC therapeutic agent for use in treatment of a pathology related to aberrant  
CC expression or physiological interactions of the polypeptide. (III) is  
CC useful for treating a pathological state in a mammal and for determining  
CC the presence or amount of POLYX in a sample. POLY1-4 (GABA receptor-like  
CC proteins) are useful for the treatment of psychiatric and medical  
CC conditions, depression, stroke, Parkinson's disease, Huntington's  
CC disease, Tourette's syndrome, amyotrophic lateral sclerosis, head trauma,  
CC Alzheimer's disease, alcoholism, vigilance, anxiety, muscle tension,  
CC epileptogenic activity and memory functions, cardiomyopathy and  
CC arrhythmogenic right ventricular dysplasia. POLY5-8 (Epidermal growth  
CC factor like proteins) may be useful for treating cancer, aberrant

CC angiogenesis, renal disease and diabetes. POLY12 (haematopoietic stem and  
CC progenitor cell like protein) may be useful for treatment of leukaemia,  
CC lupus and anaemia. POLY13 (sulforanferase-like protein) may be useful  
CC for treating cirrhosis, cholangitis, hepatitis, ulcers, hyperthyroidism  
CC and developmental disorders. POLY14-16 (Syntaxin-like proteins) may be  
CC useful in treatment of Lambert-Eaton myasthenic syndrome, asthma, myxoid  
CC liposarcoma and acute myeloid leukaemia, and POLY 18 may be useful in  
CC treatment of cancers. Cells comprising (I) are useful for producing non-  
CC human transgenic animals which are useful for studying the function  
CC and/or activity of POLYX protein and for identifying and/or evaluating  
CC modulators of POLYX protein activity. This is the amino acid sequence of  
CC a novel human protein

XX  
SQ Sequence 305 AA;

Query Match 26.1%; Score 399.5; DB 6; Length 305;

Best Local Similarity 30.9%; Pred. No. 8.5e-33;

Matches 93; Conservative 61; Mismatches 126; Indels 21; Gaps 5;

QY 1 MAESEAETPTSGEFE-SKYFEFGVRLPPFCRGKMBEIANFPVRPSDVWIVTPKSGTS 59

DB 1 MAKTEKNAPTMEKKPELFNIMEVDGVTLLILSKWEKVCNFQAKPDDLLILATPKSGTT 60

QY 60 LLOEVVYLVSGQADPDEIGLNMIDQLPVLE--YP---QPGLDIKELTSPRLIKSHLPY 114

DB 61 WMHEILDMLNDGDEVKCKRAQTLDRHAFLEKLPKHEKPDLEFVLEMSSPQLIKTHLPS 120

QY 115 RFLPSDLHNGDSKVIYMARNPCKDLVSVYQPHRSRLTMSYRGTFQFCRRFMNDKLGYS 174

DB 121 HLIPPSIWKENCKIVYVARNPKDCLVSYVYHFRMASFMPDPQNLLEEFYEKFMGKGERGS 180

QY 175 WFEHVQFWEHRMDSNVLFKXYEDMHRDLVTMVEQLARFLVGSCKDAQLEALTECHOLV 234

DB 181 WFDHVKGWMAAKDHRILYLFYEDIKQNPKEIHKVLEFLEKTSVGDVINKIVHHTSPDV 240

QY 235 ---DQCCNAEALPV-----GGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLT 279

DB 241 MKDNPMAHNTAVPAHIFNHSISKFMKCGMPGDKWKNHFTVALNENFDKHYEKKWAGSTLN 300

QY 280 F 280

DB 301 F 301

RESULT 71

AAE05178

XX AAE05178 standard; protein; 304 AA.

AC AAE05178;

XX

DT 12-SEP-2001 (first entry)

XX

XX Human drug metabolising enzyme (DME-9) protein.

DE

XX

XX Human; drug metabolising enzyme; DME-9; immunosuppressive; gene therapy;

KW cystostatic; autoimmune disorder; inflammatory disorder; atherosclerosis;

KW osteoporosis; eye disorder; hepatic tumour; Addison's disease; cretinism;

KW rheumatoid arthritis; acquired immune deficiency syndrome; AIDS; anaemia;

KW developmental disorder; endocrine disorder; iritis; acromegaly; epilepsy;

KW thyrotoxicosis; pancreatic disorder; diabetes mellitus; obesity; glaucoma;

KW gastrointestinal disorder; nodular hyperplasia; conjunctivitis; glaucoma;

KW acinic keratosis; metabolic disorder; dysphagia; anorexia; carcinoma;

XX cell proliferative disorder.

OS Homo sapiens.

XX

XX Key Location/Qualifiers

XX Binding-site 20..304

XX /label= PAPS\_cofactor\_binding\_site.

XX WO2000151638-A2.

XX 19-JUL-2001.

PD

XX 12-JAN-2001; 2001WO-US001174.

XX 14-JAN-2000; 2000US-0176139P.

PR 21-JAN-2000; 2000US-0177443P.

PR 28-JAN-2000; 2000US-0178574P.

XX (INCY-) INCYTE GENOMICS INC.

XX Yang J, Baughn MR, Burford N, Au-Young J, Lu DAM, Reddy R;

PI Ring HZ, Hillman JL, Yue H, Azimzai Y, Yao MG, Gandhi AR;

PI Nguyen DB, Tang YT, Lal P, Bandman O;

XX WPI; 2001-425874/45.

DR N-PSDB; AAD09944.

XX Drug metabolizing enzymes and encoding polynucleotides, useful for

PT diagnosing, treating and/or preventing autoimmune, inflammatory, cell

PT proliferative, developmental, endocrine, eye, metabolic, and

PT gastrointestinal disorders.

XX

PS Claim 1; Page 143-144; 133pp; English.

XX The present sequence is human drug metabolising enzyme (DME-9) protein.

CC Human DME and its nucleic acid molecule are useful for the diagnosis,

CC treatment and prevention of disorders associated with increased or

CC decreased expression of DME. Examples of such disorders include,

CC autoimmune/inflammatory disorder such as acquired immune deficiency

CC syndrome (AIDS), rheumatoid arthritis, osteoporosis; cell proliferative

CC disorder such as actinic keratosis, atherosclerosis; developmental

CC disorder such as epilepsy, anaemia; endocrine disorder such as

CC acromegaly, cretinism, thyrotoxicosis; pancreatic disorder such as

CC diabetes mellitus; eye disorder such as conjunctivitis, glaucoma, iritis;

CC metabolic disorder such as Addison's disease, obesity; gastrointestinal

CC disorder such as anorexia, dysphagia and hepatic tumours including

CC nodular hyperplasia, adenomas and carcinomas. DME DNA is useful for

CC creating 'knockin' humanised animals (pigs) or transgenic animals (mice

CC or rats) to model human disease. DME DNA is also in useful is gene

CC therapy. DME and its immunogenic fragments are useful for screening

CC libraries of compounds in several drug screening assays

XX

SQ Sequence 304 AA;

Query Match 26.1%; Score 399; DB 4; Length 304;

Best Local Similarity 31.0%; Pred. No. 9.5e-33;

Matches 93; Conservative 61; Mismatches 126; Indels 20; Gaps 5;

QY 1 MAESEAETPTSGEFE-SKYFEFGVRLPPFCRGKMBEIANFPVRPSDVWIVTPKSGTS 59

DB 1 MAKTEKNAPTMEKKPELFNIMEVDGVTLLILSKWEKVCNFQAKPDDLLILATPKSGTT 60

QY 60 LLOEVVYLVSGQADPDEIGLNMIDQLPVLE--YP---QPGLDIKELTSPRLIKSHLPY 114

DB 61 WMHEILDMLNDGDEVKCKRAQTLDRHAFLEKLPKHEKPDLEFVLEMSSPQLIKTHLPS 120

QY 115 RFLPSDLHNGDSKVIYMARNPCKDLVSVYQPHRSRLTMSYRGTFQFCRRFMNDKLGYS 174

DB 121 HLIPPSIWKENCKIVYVARNPKDCLVSYVYHFRMASFMPDPQNLLEEFYEKFMGKGVRS 180

QY 175 WFEHVQFWEHRMDSNVLFKXYEDMHRDLVTMVEQLARFLVGSCKDAQLEALTECHOLV 234

DB 181 WFDHVKGWMAAKDHRILYLFYEDIKQNPKEIHKVLEFLEKTSVGDVINKIVHHTSPDV 240

QY 235 ---DQCCNAEALPV-----GGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLT 280

DB 241 MKDNPMAHNTAVPAHIFNHSISKFMKCGMPGDKWKNHFTVALNENFDKHYEKKWAGSTLN 300

RESULT 72

ADD12583

ID ADD12583 standard; protein; 261 AA.

XX

AC ADD12583;

XX	01-JAN-2004	(first entry)	
XX	Human ENZM-43	protein SEQ ID NO:43.	
DE	human;	enzyme; ENZM; cytototoxic; antiarteriosclerotic; anti-HIV;	
XX		antiallergic; cerebroprotective; antiparkinsonian; anticonvulsant;	
KW		nootropic; neuroprotective; antiinflammatory; ophthalmological;	
KW		antithyroid; antiarthritic; antibacterial; virucide; protozoacide;	
KW		antiparasitic; fungicide; anorectic; cardiant; hypotensive;	
KW		antinfertility; hepatotropic; gene therapy; autoimmune disorder;	
KW		inflammatory disorder; AIDS; allergy; atopic dermatitis; arthritis;	
KW		thyroiditis; infection; metabolic disorder; obesity;	
KW		reproductive disorder; infertility; neurological disorder;	
KW		Parkinson's disease; Alzheimer's disease; stroke; epilepsy;	
KW		cardiovascular disorder; myocardial infarction; hypertension;	
KW		eye disorder; cell proliferative disease; cancer; atherosclerosis;	
KW		hepatitis.	
XX			
OS	Homo sapiens.		
XX			
PN	WO2003072729-A2.		
XX			
PD	04-SEP-2003.		
XX			
PF	21-FEB-2003;	2003WO-US005478.	
XX			
XX	22-FEB-2002;	2002US-0359513P.	
PR	19-MAR-2002;	2002US-0365795P.	
XX			
PA	(INCY-)	INCYTE GENOMICS INC.	
XX			
PI	Sprague WW,	Kable AE, Lee S, Griffin JA, Becha SD, Lee EA;	
PI	Hafalia AJA,	Khare R, Emerling BM, Marquis JP, Rankumar J;	
PI	Elliott VS,	Richardson TW, Baughn MR, Jin P, Chien D, Hawkins PR;	
PI	Wilson AD,	Chawla NK, Tran UK, Lee SY, Zebardjian Y, Jiang X;	
PI	Jackson AA,	Bhatia UG, Burrill JD, Blake JJ, Ho A, Zheng W;	
XX			
DR	WPI;	2003-731608/69.	
DR	N-PSDB;	ADD12640.	
XX			
PT	New human enzymes	and polynucleotides, useful for diagnosing, preventing	
PT	or treating diseases	or conditions associated with aberrant enzyme	
PT	expression, e.g.	cancer, hepatitis, AIDS, atherosclerosis, obesity,	
PT	infection or stroke.		
XX			
PS	Claim 1;	SEQ ID NO 43; 426pp; English.	
XX			
CC	The present sequence	represents a human enzyme designated ENZM-43. Human	
CC	enzymes of the present	invention have cytototoxic, antiarteriosclerotic,	
CC	anti-HIV, antiallergic,	cerebroprotective, antiparkinsonian,	
CC	anticonvulsant,	nootropic, neuroprotective, antiinflammatory,	
CC	ophthalmological,	antithyroid, antiarthritic, antibacterial, virucide,	
CC	protozoacide,	antiparasitic, fungicide, anorectic, cardiant, hypotensive,	
CC	antinfertility	and hepatotropic activities, and can be used in gene	
CC	therapy. The human	enzymes and polynucleotides encoding them can be used	
CC	in diagnosing,	preventing or treating diseases or conditions associated	
CC	with the decreased	expression or overexpression of the enzymes, such as	
CC	autoimmune/inflammatory	disorders (AIDS, allergies, atopic dermatitis,	
CC	arthritis or thyroiditis),	infections (e.g. bacterial, viral, parasitic,	
CC	protozoal or fungal),	metabolic disorders (e.g. obesity), reproductive	
CC	disorders (e.g. infertility),	neurological disorders (Parkinson's	
CC	disease, Alzheimer's	disease, stroke, epilepsy), cardiovascular disorders	
CC	(e.g. myocardial	infarction or hypertension), eye disorders, or cell	
CC	proliferative diseases	(e.g. cancer, atherosclerosis, hepatitis). They	
CC	are also useful	in assessing the effects of exogenous compounds on the	
CC	expression of nucleic	acid and amino acid sequences of the enzymes.	
XX			
SQ	Sequence 261	AA;	
	Query Match	26.0%;	Score 397.5; DB 7; Length 261;
	Best Local Similarity	30.4%;	Pred. No. 1.1e-32;
	Matches 85;	Conservative 53;	Mismatches 91; Indels 51; Gaps 3;





PR 15-MAR-2001; 2001US-0276450P.  
PR 16-MAR-2001; 2001US-0276397P.  
PR 16-MAR-2001; 2001US-0276768P.  
PR 20-MAR-2001; 2001US-0278652P.  
PR 26-MAR-2001; 2001US-0278775P.  
PR 26-MAR-2001; 2001US-0278775P.  
PR 29-MAR-2001; 2001US-0279882P.  
PR 29-MAR-2001; 2001US-0279884P.  
PR 30-MAR-2001; 2001US-0280147P.  
PR 11-APR-2001; 2001US-0282992P.  
PR 11-APR-2001; 2001US-0283083P.  
PR 20-APR-2001; 2001US-0285133P.  
PR 23-APR-2001; 2001US-0285743P.  
PR 03-MAY-2001; 2001US-0288327P.  
PR 03-MAY-2001; 2001US-0288504P.  
PR 29-MAY-2001; 2001US-0294047P.  
PR 30-MAY-2001; 2001US-0294473P.  
PR 08-JUN-2001; 2001US-0296964P.  
PR 18-JUN-2001; 2001US-0298959P.  
PR 19-JUN-2001; 2001US-0299324P.  
PR 13-AUG-2001; 2001US-0312020P.  
PR 16-AUG-2001; 2001US-0312889P.  
PR 16-AUG-2001; 2001US-0312908P.  
PR 21-AUG-2001; 2001US-0313930P.  
PR 28-AUG-2001; 2001US-0315470P.  
PR 31-AUG-2001; 2001US-0316447P.  
PR 07-SEP-2001; 2001US-0318115P.  
PR 07-SEP-2001; 2001US-0318118P.  
PR 12-SEP-2001; 2001US-0318740P.  
PR 19-SEP-2001; 2001US-0323379P.  
PR 18-OCT-2001; 2001US-0330245P.  
PR 18-OCT-2001; 2001US-0330308P.  
PR 14-NOV-2001; 2001US-0332701P.  
PA (TCHE/) TCHERNEV V T.  
PA (SPYK/) SPYTEK K A.  
PA (ZERH/) ZERHUSEN B D.  
PA (PATI/) PATTURAJAN M.  
PA (SHIM/) SHIMKETS R A.  
PA (LILL/) LI L.  
PA (GANG/) GANGOLLI E A.  
PA (PADI/) PADIGARU M.  
PA (ANDE/) ANDERSON D W.  
PA (RAST/) RASTELLI L.  
PA (MILL/) MILLER C E.  
PA (GERL/) GERLACH V.  
PA (TAUP/) TAUPIER R J.  
PA (GUSE/) GUSEV V Y.  
PA (COLM/) COLMAN S D.  
PA (WOLE/) WOLENC A R.  
PA (PENA/) PENNA C E A.  
PA (FURT/) FURTAK K.  
PA (GROS/) GROSSE W M.  
PA (ALSO/) ALSOBROOK J P.  
PA (LEPL/) LEPLEY D M.  
PA (RIEG/) RIEGER D K.  
PA (BURG/) BURGESS C E.  
XX  
PI Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;  
PI Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;  
PI Gerlach V, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CE;  
PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;  
XX WPI; 2004-180039/17.  
DR N-PSDB; ADN42323.  
XX  
XX Isolated NOVX polypeptides and polynucleotides, useful for preventing  
XX diagnosing and/or treating cancer, diabetes, cardiomyopathy and  
XX atherosclerosis.  
PS Claim 1; SEQ ID NO 206; 1309pp; English.  
XX  
XX The invention relates isolated 162 NOVX polypeptides (NOV1-NOV99,  
CC

CC including splice variants) and the nucleic acids (NA) that encode them.  
CC Also included are the mature NOVX proteins (and their encoding the  
CC polynucleotides), a vector comprising NOVX NA, a cell comprising the  
CC vector, an antibody that binds immunospecifically to NOVX, determining  
CC the presence or amount of NOVX in a sample, determining the presence or  
CC amount of NOVX NA in a sample, identifying an agent that binds to NOVX,  
CC modulating the activity of NOVX, treating or preventing a NOVX-associated  
CC disorder, determining the presence of or predisposition to a disease  
CC associated with altered levels of NOVX and treating a pathological state  
CC in a mammal comprising administering a polypeptide which is at least 95%  
CC identical to NOVX (or fragment). NOVX and NA may be used in the  
CC prevention, treatment and diagnosis of diseases associated with  
CC inappropriate expression and activity of NOVX (e.g. cancer, diabetes,  
CC cardiomyopathy and/or atherosclerosis). The anti-NOVX antibodies and  
CC antagonists may also be used to down regulate expression and activity of  
CC NOVX. The anti-NOVX antibodies may also be used as diagnostic agents for  
CC detecting the presence of NOVX in samples (e.g. by enzyme linked  
CC immunosorbant assay (ELISA). The agents and methods may be used in this  
CC way to prevent, diagnose and treat cancer, diabetes, cardiomyopathy  
CC and/or atherosclerosis. The present sequence represents a NOVX protein.  
XX  
SQ Sequence 295 AA;  
Query Match 26.0%; Score 397.5; DB 8; Length 295;  
Best Local Similarity 32.2%; Pred. No. 1.3e-32;  
Matches 85; Conservative 58; Mismatches 102; Indels 19; Gaps 4;  
QY 36 EEIANFPVRPSDVIWVTPKSGTSLLOEVYLVSGADPDDEIGLWNIDEQLPVLE--Yp- 92  
DB 28 EKVCNFOAKPDDLLIATYPKSGTTWHEILDMILNDGDVEKCKRAQTLDRHAFLELKEPH 87  
QY 93 --QPLDIIKELTSPRLIKGHLPYRFLPSDLHNGDSKVIYMARPKDLVSVYFPHSLR 150  
DB 88 KEKPDLEFVLEMSPPQLIKTHLPShLIPPSIWKENCIVYVARNPKDCLVSVYHFRMAS 147  
QY 151 TMSYRGTFQRCRPFMNDKLGYSWFHVOEFWEHRMDSNVLFKYEYDMHRDLVTWVQL 210  
DB 148 FMPDPQNLBEFYEKFGSKVGVGWFDMKGMWAAKDMHRLILYLFYEDIKKPKHETHKV 207  
QY 211 ARFLGVSCKRAQLEALTEHCHQLV---DQCNAEALP-----VGRHVGVLWKDI 256  
DB 208 LEFLEKTSWGDVINKIVHHTSFVWKNPNMANHTAVPAHIFNHISIKFMKGMFGDWKQH 267  
QY 257 FTVSMNEKFDLVYKQKMGKCDLTF 280  
DB 268 FTVAMNENFDKHEKMGAGSTLNF 291  
RESULT 75  
ID AAG68266 standard; protein; 305 AA.  
XX AAG68266;  
XX  
XX 14-FEB-2002 (first entry)  
XX Human POLY13 protein sequence SEQ ID NO:26.  
XX  
XX Human; POLYX; gamma aminobutyric acid receptor; GABA receptor;  
KW Human; POLYX; gamma aminobutyric acid receptor; GABA receptor;  
KW haematopoietic stem and progenitor cell; sulphotransferase; prohibitin;  
KW antidepressant; neuroprotective; antiparkinsonian; nootropic; relaxant;  
KW anticonvulsant; neuroleptic; neuroprotective; antialcoholic; cardiant;  
KW tranquiliser; antiarrhythmic; psychiatric; medical; depression; stroke;  
KW Parkinson's disease; Huntington's disease; Tourette's disease;  
KW amyotrophic lateral sclerosis; head trauma; Alzheimer's disease;  
KW alcoholism; vigilance; muscle tension; epileptogenic; memory function;  
KW cardiomyopathy; arrhythmogenic right ventricular dysplasia; chromosome 2.  
XX  
XX Homo sapiens.  
XX  
XX W0200179294-A2.  
XX

PD 25-OCT-2001.

XX 19-APR-2001; 2001WO-US012854.

PF 19-APR-2000; 2000US-0198293P.

XX 20-APR-2000; 2000US-0198645P.

PR 25-APR-2000; 2000US-0199476P.

PR 26-APR-2000; 2000US-0199880P.

PR 26-APR-2000; 2000US-0200024P.

PR 26-APR-2000; 2000US-0200025P.

PR 09-JUN-2000; 2000US-0210809P.

PR 17-JUL-2000; 2000US-0218591P.

PR 11-AUG-2000; 2000US-0224610P.

PR 09-FEB-2001; 2001US-0267673P.

PR 27-FEB-2001; 2001US-0271814P.

XX (CURA-) CURAGEN CORP.

XX Taupier RJ, Vernet CAM, Fernandes E, Shimkets RA, Majumder K;

PI Padigaru M, Colman SD, Spytek KA, Burgess CE, Liu X;

PI N-PSDB; ABA03882.

DR WPI; 2002-017601/02.

DR New isolated polypeptides for treating a broad range of pathological

XX states, e.g., depression, stroke, Parkinson's disease, Huntington's

PT disease, Tourette's syndrome, amyotrophic lateral sclerosis, head trauma,

PT and Alzheimer's.

XX Claim 1; Page 52-53; 155pp; English.

XX The present invention describes polypeptides (I), designated POLYX

CC polypeptides (e.g. POLY1, POLY2, etc.), and the polynucleotide sequences

CC (II) encoding them. POLY1-4 are members of the gamma aminobutyric acid

CC (GABA) receptor family; POLY5-8 are members of the epidermal growth

CC factor (EGF) family; POLY9-11 are members of the complement receptor

CC family; POLY12 is a member of the haematopoietic stem and progenitor cell

CC (HSPC) family; POLY13 is a member of the sulphotransferase family; POLY14

CC -16 are members of the syntaxin family; and POLY17 is a member of the

CC prohibitin family. (I) and (II) can have antidepressant,

CC cerebrotective, antiparkinsonian, nootropic, relaxant, anticonvulsant,

CC neuroleptic, neuroprotective, antialcoholic, radiant, tranquilliser and

CC antiarrhythmic activities. (I) and (II) can be used for treating or

CC preventing a POLYX-associated disorder in humans as a therapeutic in the

CC manufacture of a medicament for treating a syndrome associated with a

CC human disease selected from a POLYX-associated disorder, for treating a

CC pathological state in a mammal, especially patients suffering from, e.g.,

CC psychiatric and medical conditions, depression, stroke, Parkinson's

CC disease, Huntington's disease, Tourette's syndrome, amyotrophic lateral

CC sclerosis, head trauma, Alzheimer's disease, alcoholism, vigilance,

CC anxiety, muscle tension, epileptogenic activity and memory functions,

CC cardiomyopathy and arrhythmogenic right ventricular dysplasia. The

CC present sequence represents POLY13, which was identified on chromosome 2

XX Sequence 305 AA;

SQ

Query Match 25.9%; Score 396.5; DB 5; Length 305;

Best Local Similarity 30.9%; Pred. No. 1.7e-32;

Matches 93; Conservative 60; Mismatches 127; Indels 21; Gaps 5;

QY 1 MAESAEPTPTGPEFE-SKYFEFHGVRLPPFCRGHMEIANFPVRPSDVWIVTPKSGTS 59

DB 1 MAKEIKNAPTWEKKPELNEINVEDGVPTLLISKEWKEVCNFOAKPDDLIIATPKSGTT 60

QY 60 LLQEVYVLVSGQADPEIGLNMIDRLQFVLE--YP---QPGLDITIKELTSPRLKSHLPY 114

DB 61 WNHETLMDILNDGDVEKCKRAQTLDRAFLFKFPHKEKPDLEFVLEMSSPQLIKTHLPS 120

QY 115 RFLPSDLNLSGSKVITYMARNPKDLVSYVQFHRSLRSTYRGTFQFCRFRWNXLGVGS 174

DB 121 HLIPPSIKWENCKIVYVARNPKDLVSYVQFHRMASFMPDPQNLBEFYKFMGSGGFGFS 180

QY 175 WFEHVQEPWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLEALTEHCHQLV 234

Db 181 WFDHVKGWAAKDMHRILYLFYEDIKQNPKEIHKVLEFLEKTWSGDVINKIVHTSPDV 240

QY 235 ---DQCCNAEALPV-----GRGRVGLWKDIFTVSMNEKFDLVYKQKVGKCDLT 279

Db 241 MKONPMANHTAVPAHIFNHSISKFMKGGMPGDWKNHFTVALNENFDKHEKMGAGSTLN 300

QY 280 F 280

Db 301 F 301

RESULT 76

AAE39814

ID AAE39814 standard; protein; 305 AA.

XX AC AAE39814;

XX 18-DEC-2003 (first entry)

XX Human sulfotransferase-like protein (POLY13).

XX Human; tumour; inflammatory disorder; vaccine; gene therapy; cytostatic;

XX sulfotransferase-like protein; POLY13; chromosome 2.

XX Homo sapiens.

XX US2003050232-A1.

XX 13-MAR-2003.

XX 19-APR-2001; 2001US-00839446.

XX 19-APR-2000; 2000US-0198293P.

XX 20-APR-2000; 2000US-0198645P.

XX 25-APR-2000; 2000US-0199476P.

XX 26-APR-2000; 2000US-0199880P.

XX 26-APR-2000; 2000US-0200024P.

XX 26-APR-2000; 2000US-0200025P.

XX 09-JUN-2000; 2000US-0210809P.

XX 17-JUL-2000; 2000US-0218591P.

XX 11-AUG-2000; 2000US-0224610P.

XX 27-FEB-2001; 2001US-0271814P.

XX (TAUP/) TAUPIER R J.

PA (PADI/) PADIGARU M.

PA (SPYT/) SPYTEK K A.

PA (BURG/) BURGESS C E.

PA (VERN/) VERNET C A M.

PA (FERN/) FERNANDES E R.

PA (SHIM/) SHIMKETS R A.

PA (LIUX/) LIU X.

PA (MAJU/) MAJUMDER K.

PA (COLM/) COLMAN S D.

PA (ZERH/) ZERHUSEN B D.

XX Taupier RJ, Padigaru M, Spytek KA, Burgess CE, Vernet CAM;

PI Fernandes ER, Shimkets RA, Liu X, Majumder K, Colman SD;

PI Zerhusen BD;

XX WPI; 2003-605764/57.

DR N-PSDB; AAD58965.

XX New POLYX nucleic acid, useful for preparing a composition for treating

PT or preventing e.g., tumor or inflammatory disorder.

XX Claim 1; Page 35; 75pp; English.

CC The invention relates to new POLYX nucleic acid useful for preparing a

CC composition for treating or preventing tumour or inflammatory disorder.

CC The invention is useful as vaccine and in gene therapy. The nucleic acid

CC is useful for preparing a composition for treating or preventing e.g.,

CC tumour or inflammatory disorder. The present sequence is human

CC sulfotransferase-like protein (POLY13). POLY13 gene is located on  
CC chromosome 2. Note: The present sequence is encoded by a DNA containing  
CC translational exceptions  
XX  
SQ Sequence 305 AA;  
  
Query Match 25.9%; Score 396.5; DB 7; Length 305;  
Best Local Similarity 30.9%; Pred. No. 1.7e-32;  
Matches 93; Conservative 60; Mismatches 127; Indels 21; Gaps 5;  
  
QY 1 MAESAEATSTGPEP-SKYFEPFHGVRLPPFCRGKMEETANFPVPDVMVITYPKSGTS 59  
DB 1 MAKIEKNAFTMEKKPELFNIMEVDGVTLLSKWEWKCVCNFOAKPDDDLILATYPSGTT 60  
  
QY 60 LQEVYLVSOQADPEIGLMNIDEQLPVLE--YP---QPGLDIIKELTSPRLIKSHLPY 114  
DB 61 WMHEILDMILNDGDVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSPPQIKTHLPS 120  
  
QY 115 RFLPSDLHNGDSKVIYMARNPDLVSVYQFHRSLRTMSYRGTFQFCRRFNDKLGYS 174  
DB 121 HLIPPSIWKENCKIVVVARNPDKCLVSYVYHFRMASFMPDPQNLBEFYEKFMGSGKEFGS 180  
  
QY 175 WFEHQVEFWEHRMDSNVLFLKYEDMHRDLVTVVEQLARFLGVCSDKAQLEALTEHCHQLV 234  
DB 181 WFDHVKGWMAAKDMHRILYLFYEDIKQNPKEIHKVLEFLEKTSWGDVINKIVHHTSFQV 240  
  
QY 235 ---DOCCNAEALPV-----GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLT 279  
DB 241 MKDNPMANHTAVPAHFNHSISKFMKKGMPGDKNHNFTVALNENFDKHYEKKMAGSTLN 300  
  
QY 280 F 280  
DB 301 F 301  
  
RESULT 77  
ADRI9665  
ID ADRI9665 standard; protein; 297 AA.  
XX  
AC ADRI9665;  
XX  
DT 07-OCT-2004 (first entry)  
XX  
DE Human drug metabolising enzyme (DME) -4 protein sequence.  
XX  
KW drug metabolising enzyme; DME; cytostatic; immunosuppressive;  
KW antinflammatory; endocrine; ophthalmological; gastrointestinal;  
KW hepatotropic; cancer; cell proliferative disorder; autoimmune disorder;  
KW inflammatory disorder; endocrine disorder; eye disorder;  
KW gastrointestinal disorder; liver disorder; metabolic disorder; enzyme;  
human.  
XX  
OS Homo sapiens.  
XX  
PN WO200226988-A2.  
XX  
PD 04-APR-2002.  
XX  
PF 28-SEP-2001; 2001WO-US030662.  
XX  
PR 29-SEP-2000; 2000US-0236947P.  
PR 06-OCT-2000; 2000US-0238864P.  
PR 20-OCT-2000; 2000US-0242323P.  
PR 09-NOV-2000; 2000US-0247581P.  
PR 16-NOV-2000; 2000US-0249519P.  
PR 22-NOV-2000; 2000US-0252834P.  
PR 30-NOV-2000; 2000US-0250567P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Azinzai Y, Baughin MR, Borowsky ML, Ding L, Duggan BM;  
PI Elliott VS, Gandhi AR, Griffin JA, Hafalia AJA, Ison CH, Khan PA;  
PI Lal P, Lee EA, Lu DAM, Nguyen DB, Arvizu C, Pollicky JL, Ramkumar J;

PI Ring HZ, Sanjanwala MS, Tang YT, Tribouley CM, Narinder WK;  
PI Walsh RT, Warren BA, Xu Y, Yang J, Yao MG, Yue H;  
XX  
DR WPI: 2002-362498/39.  
DR N-PSDB; ADRI9683.  
XX  
PT Human drug metabolizing enzymes, useful in the diagnosis and treatment of  
PT disorders associated with aberrant (DME) activity, e.g., cancer and  
PT autoimmune disorders.  
XX  
PS Claim 1; SEQ ID NO 4; 142pp; English.  
XX  
CC This invention relates to novel drug metabolising enzymes (DME) and the  
CC nucleotide sequences which encode them. The invention may be useful for  
CC the development of compounds with a cytostatic, immunosuppressive,  
CC antinflammatory, endocrine, ophthalmological, gastrointestinal or  
CC hepatotropic activity acting as an agonist or antagonist of drug  
CC metabolising enzyme activity. The invention may be used in the diagnosis  
CC and treatment of disorders associated with decreased or increased  
CC expression or activity of drug metabolising enzymes. Such disorders  
CC include cancer, cell proliferative disorders, autoimmune/inflammatory,  
CC endocrine, eye, gastrointestinal (including liver disorders) and  
CC metabolic disorders. The present sequence is that of a human drug  
CC metabolising enzyme (DME) of the invention. Note: This sequence did not  
CC form part of the printed specification but was obtained in electronic  
CC format from EPO.  
XX  
SQ Sequence 297 AA;  
  
Query Match 25.2%; Score 385.5; DB 5; Length 297;  
Best Local Similarity 32.4%; Pred. No. 2.4e-31;  
Matches 84; Conservative 55; Mismatches 105; Indels 15; Gaps 3;  
  
QY 36 BEIANFPVRDVMIVTPYKSGTSLQEVYLVSOQADPEIGLMNIDEQLPVLEVPOPG 95  
DB 36 DKINWFOAKPDDLLISTYTPKAGTTWMHEILDMILNDGDVEKCKRAQTLDRHAFLELKP- 94  
  
QY 96 LDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDLVSVYQFHRSLRTMSYR 155  
DB 95 LEFVLEMSPPQIKTHLPSHLIPPSIWKENCKIVVVARNPDKCLVSYVYHFRMASFMPDP 154  
  
QY 156 GTFQFCRRFNDKLGYSWFEHQVEFWEHRMDSNVLFLKYEDMHRDLVTVVEQLARFLG 215  
DB 155 QNLEEFYEKFMGSGKVGVSFDFHVKGWMAAKDTHRILYLFYEDIKKNPKHEIKVLEFLE 214  
  
QY 216 VSCDKAOLEALTEHCHQLV---DOCCNAEALP-----VGRGRVGLWKDIFTVSM 261  
DB 215 KTLGSDVINKIVHHTSFQVMDKNPMANHTAVPAHFNHSISKFMKKGMPGDKNHNFTVAM 274  
  
QY 262 NEKFDLVYKQKMGKCDLTF 280  
DB 275 NENFDKHYEKKMAGSTLNF 293  
  
RESULT 78  
ABM84008  
ID ABM84008 standard; protein; 262 AA.  
XX  
AC ABM84008;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human diagnostic and therapeutic pprotein SEQ ID NO:4257.  
KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.  
XX  
OS Homo sapiens.  
XX  
PN WO2004023973-A2.  
XX  
PD 25-MAR-2004.  
XX  
PF 12-SEP-2003; 2003WO-US028227.

```
XX 12-SEP-2002; 2002US-0410259P.
PR 12-SEP-2002; 2002US-0410260P.
XX (INCY-) INCYTE CORP.
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.
DR N-PSDB; ACN42660.
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX in gene mapping.
PS Claim 27; Page; 190pp; English.
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX Sequence 262 AA;
Query Match 25.1%; Score 384.5; DB 8; Length 262;
Best Local Similarity 30.5%; Pred. No. 2.6e-31;
Matches 84; Conservative 49; Mismatches 91; Indels 51; Gaps 4;
QY 24 GVRLPFCRCRGMEEIANFPVRPSDVWIVITPKSGTSLLOEVYLVVSQGDPEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDMIYQGGDLEKCNRAPIY 76
QY 84 BOLPVLEYPPQGLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYVARNPKDLVWSY 143
DB 77 VRVPLEVNDP-----GEPGSGVYVARNPKDVAVSY 107
QY 144 QFHRSLRTMSYRGTFQEFRCRPFMDKLGYSWFEHVFQEFWHRMDSNVLFKLYEDMHRDL 203
DB 108 HFHRMEKAHPPEGTWDSFLEKFMAGEVSGYSYQHVQEWELSRTHPYLYLFYEDMKENP 167
QY 204 VTMVQLARFLGVSCDKAQLAETHC-----HQLVDQCCNAEALP-V 245
DB 168 KREIKIIEFVGRSIPETMDPMVQHTSFENKKNPMNTYTVPQELMDHSIS-----PPM 223
QY 246 GRGRVGLMKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 224 RKGWAGDNKTTFTVAQNERFDADYAEKMGACSLSF 258
RESULT 79
ABM82616
ID ABM82616 standard; protein; 262 AA.
XX
```

```
AC ABM82616;
XX 18-NOV-2004 (first entry)
XX Human diagnostic and therapeutic pprotein SEQ ID NO:2865.
XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX Homo sapiens.
XX WO2004023973-A2.
XX 25-MAR-2004.
XX 12-SEP-2003; 2003WO-US028227.
XX 12-SEP-2002; 2002US-0410259P.
PR 12-SEP-2002; 2002US-0410260P.
XX (INCY-) INCYTE CORP.
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.
DR N-PSDB; ACN41268.
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX in gene mapping.
PS Claim 27; Page; 190pp; English.
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX Sequence 262 AA;
Query Match 25.1%; Score 384.5; DB 8; Length 262;
Best Local Similarity 30.5%; Pred. No. 2.6e-31;
Matches 84; Conservative 49; Mismatches 91; Indels 51; Gaps 4;
QY 24 GVRLPFCRCRGMEEIANFPVRPSDVWIVITPKSGTSLLOEVYLVVSQGDPEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDMIYQGGDLEKCNRAPIY 76
QY 84 BOLPVLEYPPQGLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYVARNPKDLVWSY 143
DB 77 VRVPLEVNDP-----GEPGSGVYVARNPKDVAVSY 107
QY 144 QFHRSLRTMSYRGTFQEFRCRPFMDKLGYSWFEHVFQEFWHRMDSNVLFKLYEDMHRDL 203
```



CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be  
CC used to diagnose a particular condition, disease or disorder associated  
CC with human molecules, e.g. cell proliferative disorders,  
CC autoimmune/inflammatory disorder, developmental disorder, endocrine  
CC disorder, neurological disorders, gastrointestinal disorders, or  
CC infections caused by virus, bacteria, fungi or parasite. The dithp  
CC molecules may also be used in genetic mapping, in identifying individuals  
CC from minute biological samples in detecting single nucleotide  
CC polymorphisms, as molecular weight markers, and for somatic or germline  
CC gene therapy. The present sequence represents a dithp protein of the  
CC invention. Note: The sequence data for this patent is not represented in  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)  
XX  
SQ Sequence 253 AA;

Query Match 25.0%; Score 383; DB 8; Length 253;  
Best Local Similarity 30.6%; Pred. No. 3.5e-31;  
Matches 83; Conservative 45; Mismatches 91; Indels 52; Gaps 3;  
QY 24 GVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSLQEVVYLVSQGADPDEIGLMNID 83  
DB 17 GVPLIKYFAEALGPLQSQFARPDLLISTYPKS----- 49  
QY 84 EQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVSY 143  
DB 50 -----GMETLKOTPAPLLKTLPLALLPOTLLDQKVVYVARNAKDVASY 98  
QY 144 QFHSLRTMSYRGTFQFCRRFMDKLGYSWFEHVFQWFEHRMDSNVFLFKYEDMRDL 203  
DB 99 HFYHMAKVHPGPGTWSFLEKFMVGEVSGYQVQSWBELLRTHPVLYLFYEDMKENP 158  
QY 204 VTMVEQLARFLGVSCDKAQLAETHC---HQLVDQCCNAEALP-----VGRGR 249  
DB 159 KREIQKILEFVGRSPTEPTVDFVQHTSFKEKMKNPMTNTYTPQEFMDHSISPFMRKGM 218  
QY 250 VGLWKDIFTVSNNKFDLVYKQMGKCDLTF 280  
DB 219 AGDWKTTTVAQNERFDADYAEKMGCSLSF 249

RESULT 82  
ADI16668  
XX ADI16668 standard; protein; 324 AA.  
AC ADI16668;  
XX  
DT 15-APR-2004 (first entry)  
XX  
DE Human NOVX protein to treat human pathological conditions SeqID204.  
XX human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;  
KW inflammation; autoimmune disorder; allergy; blood disorder;  
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;  
KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;  
KW Alzheimer's disease; infection; stroke; muscular dystrophy; epilepsy;  
KW cysticatic; cardiac; antiinflammatory; immunosuppressive; antiallergic;  
KW haemostatic; anti-HIV; antidiabetic; antiarteriosclerotic; anorectic;  
KW antiasthmatic; nephrotropic; antiarthritic; hepatotropic;  
KW neuroprotective; nootropic; antibacterial; virucide; antiparasitic;  
KW relaxant; anticonvulsant; neurogenesis; wound healing; angiogenesis;  
KW chromosome mapping; tissue typing; pharmacogenomic; SNP;  
KW single nucleotide polymorphism.  
XX  
OS Homo sapiens.  
XX  
XX WO200268649-A2.  
PN  
XX  
PD 06-SEP-2002.  
XX  
PF 31-JAN-2002; 2002WO-US002785.  
XX  
XX 31-JAN-2001; 2001US-0265395P.  
PR

PR 31-JAN-2001; 2001US-0265412P.  
PR 31-JAN-2001; 2001US-0265514P.  
PR 31-JAN-2001; 2001US-0265517P.  
PR 02-FEB-2001; 2001US-0266406P.  
PR 05-FEB-2001; 2001US-0266767P.  
PR 07-FEB-2001; 2001US-0266975P.  
PR 07-FEB-2001; 2001US-0267057P.  
PR 08-FEB-2001; 2001US-0267459P.  
PR 09-FEB-2001; 2001US-0267823P.  
PR 15-FEB-2001; 2001US-0268974P.  
PR 26-FEB-2001; 2001US-0271664P.  
PR 27-FEB-2001; 2001US-0271839P.  
PR 27-FEB-2001; 2001US-0271855P.  
PR 02-MAR-2001; 2001US-0272788P.  
PR 02-MAR-2001; 2001US-0273046P.  
PR 14-MAR-2001; 2001US-0275925P.  
PR 14-MAR-2001; 2001US-0275947P.  
PR 14-MAR-2001; 2001US-0275950P.  
PR 14-MAR-2001; 2001US-0275989P.  
PR 15-MAR-2001; 2001US-0276448P.  
PR 15-MAR-2001; 2001US-0276450P.  
PR 16-MAR-2001; 2001US-0276397P.  
PR 16-MAR-2001; 2001US-0276768P.  
PR 20-MAR-2001; 2001US-0278652P.  
PR 26-MAR-2001; 2001US-0278775P.  
PR 26-MAR-2001; 2001US-0278778P.  
PR 29-MAR-2001; 2001US-0279882P.  
PR 29-MAR-2001; 2001US-0279884P.  
PR 30-MAR-2001; 2001US-0280147P.  
PR 11-APR-2001; 2001US-0282992P.  
PR 11-APR-2001; 2001US-0283083P.  
PR 20-APR-2001; 2001US-0285133P.  
PR 23-APR-2001; 2001US-0285749P.  
PR 03-MAY-2001; 2001US-0288327P.  
PR 03-MAY-2001; 2001US-0288504P.  
PR 29-MAY-2001; 2001US-0294047P.  
PR 30-MAY-2001; 2001US-0294473P.  
PR 08-JUN-2001; 2001US-0296964P.  
PR 18-JUN-2001; 2001US-0298959P.  
PR 19-JUN-2001; 2001US-0299344P.  
PR 13-AUG-2001; 2001US-0312889P.  
PR 16-AUG-2001; 2001US-0312908P.  
PR 21-AUG-2001; 2001US-0313390P.  
PR 28-AUG-2001; 2001US-0315470P.  
PR 31-AUG-2001; 2001US-0316447P.  
PR 07-SEP-2001; 2001US-0318115P.  
PR 07-SEP-2001; 2001US-0318118P.  
PR 12-SEP-2001; 2001US-0318740P.  
PR 19-SEP-2001; 2001US-0323379P.  
PR 18-OCT-2001; 2001US-0330245P.  
PR 18-OCT-2001; 2001US-0330308P.  
PR 14-NOV-2001; 2001US-0332701P.  
PA (CURA-) CURAGEN CORP.  
XX Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;  
PI Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;  
PI Gerlach VL, Taupier RU, Gusev VY, Colman SD, Wolenc AR, Pena CE;  
PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Burgess CE;  
XX WPI; 2002-706998/76.  
DR N-PSDB; ADI16667.  
XX  
XX New NOVX polypeptides and nucleic acids, useful for preventing or  
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,  
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX  
PS Claim 1; SEQ ID NO 204; 1498pp; English.  
XX  
XX This invention relates to a novel nucleic acids, and encoded polypeptides  
CC thereof, which have properties related to the stimulation of biochemical



PR 16-AUG-2001; 2001US-0312908P.  
 PR 21-AUG-2001; 2001US-0313930P.  
 PR 28-AUG-2001; 2001US-0315470P.  
 PR 31-AUG-2001; 2001US-0316447P.  
 PR 07-SEP-2001; 2001US-0318115P.  
 PR 07-SEP-2001; 2001US-0318115P.  
 PR 12-SEP-2001; 2001US-0318740P.  
 PR 19-SEP-2001; 2001US-0323379P.  
 PR 18-OCT-2001; 2001US-0330245P.  
 PR 18-OCT-2001; 2001US-0330308P.  
 PR 14-NOV-2001; 2001US-0332701P.  
 XX (TCHE/) TCHERNEV V T.  
 PA (SPVT/) SPYTEK K A.  
 PA (ZERH/) ZERHUSEN B D.  
 PA (PATT/) PATTURAJAN M.  
 PA (SHIM/) SHIMKETS R A.  
 PA (LILL/) LI L.  
 PA (GANG/) GANGOLLI E A.  
 PA (PADI/) PADIGARU M.  
 PA (ANDE/) ANDERSON D W.  
 PA (RAST/) RASTELLI L.  
 PA (MILL/) MILLER C E.  
 PA (GERL/) GERLACH V.  
 PA (TAUP/) TAUPIER R J.  
 PA (GUSE/) GUSEV V V.  
 PA (COLM/) COLMAN S D.  
 PA (MOLE/) MOLENC A R.  
 PA (PENA/) PENNA C E A.  
 PA (FURT/) FURTA K.  
 PA (GROS/) GROSSE W M.  
 PA (ALSO/) ALSOBROOK J P.  
 PA (LEPL/) LEFLEY D M.  
 PA (RIEG/) RIEGER D K.  
 PA (BURG/) BURGESS C E.  
 XX  
 PI Tchernev VT, Spyttek KA, Zerhusen BD, Patturajan M, Shimkets RA;  
 PI Li L, Gangolli EA, Padigar M, Anderson DW, Rastelli L, Miller CE;  
 PI Gerlach V, Taupier RJ, Gusev VV, Colman SM, Wolenc AR, Penna CE;  
 PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;  
 XX  
 DR WPI; 2004-180039/17.  
 DR N-PSDB; ADN42321.  
 XX  
 PT Isolated NOVX polypeptides and polynucleotides, useful for preventing  
 PT diagnosing and/or treating cancer, diabetes, cardiomyopathy and  
 PT atherosclerosis.  
 XX  
 PS Claim 1; SEQ ID NO 204; 1309pp; English.  
 XX  
 CC The invention relates isolated 162 NOVX polypeptides (NOV1-NOV99,  
 CC including splice variants) and the nucleic acids (NA) that encode them.  
 CC Also included are the mature NOVX proteins (and their encoding  
 CC polynucleotides), a vector comprising NOVX NA, a cell comprising the  
 CC vector, an antibody that binds immunospecifically to NOVX, determining  
 CC the presence or amount of NOVX in a sample, determining the presence or  
 CC amount of NOVX NA in a sample, identifying an agent that binds to NOVX,  
 CC modulating the activity of NOVX, treating or preventing a NOVX-associated  
 CC disorder, determining the presence of or predisposition to a disease  
 CC associated with altered levels of NOVX and treating a pathological state  
 CC in a mammal comprising administering a polypeptide which is at least 95%  
 CC identical to NOVX (or fragment). NOVX and NA may be used in the  
 CC prevention, treatment and diagnosis of diseases associated with  
 CC inappropriate expression and activity of NOVX (e.g. cancer, diabetes,  
 CC cardiomyopathy and/or atherosclerosis). The anti-NOVX antibodies and  
 CC antagonists may also be used to down regulate expression and activity of  
 CC NOVX. The anti-NOVX antibodies may also be used as diagnostic agents for  
 CC detecting the presence of NOVX in samples (e.g. by enzyme linked  
 CC immunosorbent assay (ELISA)). The agents and methods may be used in this  
 CC way to prevent, diagnose and treat cancer, diabetes, cardiomyopathy  
 CC and/or atherosclerosis. The present sequence represents a NOVX protein.  
 XX  
 SQ Sequence 324 AA;

Query Match 25.0%; Score 183; DB 8; Length 324;  
 Best Local Similarity 28.8%; Pred. No. 4.9e-31;  
 Matches 92; Conservative 62; Mismatches 126; Indels 40; Gaps 6;  
 QY 1 MAESEATPSPGFE-SKYPEFHGVRLLPPPCRGMBEIANFPVRPSDV----- 48  
 DB 1 MAKIEKNAPTMEKPELFNIMEVDGVTLLILSKWEKVCNFOAKPDDLIHASMLYLTIG 60  
 QY 49 -----WIVTYPKGTSLLQENVVILVSGADPDEIGLMNIDEQLPVLE--YP---OP 94  
 DB 61 KLPEEDHQAWLGNYPKSGTTWMHEILDMLNDGVEKCKRAQTLDRAFLKLPFKPEKP 120  
 QY 95 GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPOLVSVYQPHRSRLTMSY 154  
 DB 121 DLEFVLEMSQPLIKTHLPShLIPPSIWKENCKIVYVTRNPKDCLVSYIHFHRWASFPMD 180  
 QY 155 RGTQFEFCRRFEMNDKLGYSWFHVBQFWEHRMDSNVLFKYEDMHRDLVTMVQLARFL 214  
 DB 181 PQNLEEFVEKFMCKVVGGSWFDHMKGWAAKDMHRIILYFYEDIKKPKHEIHKVLEFL 240  
 QY 215 GVSCDKAQLEALTEHCHQLV---DQCNAEALP-----VGRGRVLWKDIFTVS 260  
 DB 241 EKTWSGDVINKIVHTSFDMKDNPMANTAVPAHIFNHSISKPMKGMPCGDKNHTVA 300  
 QY 261 MNEKFDLVYKQKMGCKDLTF 280  
 DB 301 MNENFDKHYEKMGAGSTLNF 320  
 RESULT 84  
 AAU18710  
 ID AAU18710 standard; protein; 76 AA.  
 XX AC AAU18710;  
 XX DT 21-NOV-2001 (first entry)  
 XX DE Renal and cardiovascular-associated protein, Seq ID 149.  
 XX KW Human; antiinflammatory; neuroprotective; immunomodulator; vulnery;  
 KW cardiovascular; cytostatic; nephrotropic; antianemic; nephritis;  
 KW immunosuppressive; kidney disorder; renal failure; hypertension;  
 KW cardiovascular disorder; myocardial infarction; blood disorder; anaemia;  
 KW blood coagulation disorder; electrolyte imbalance disorder; cancer;  
 KW hyponatraemia; hyperkalaemia; neoplastic disorder; nephroma;  
 KW autoimmune disease; inflammatory disease; reproductive system disorder;  
 KW endocrine disorder; neural activity; neurological disorder;  
 KW wound healing; respiratory disorder.  
 XX OS Homo sapiens.  
 XX PN WO200155328-A2.  
 XX PD 02-AUG-2001.  
 XX PF 17-JAN-2001; 2001WO-US001359.  
 XX PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.



PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225470P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 21-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241224P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241828P.  
PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 01-DEC-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0251990P.  
PR 05-JAN-2001; 2001US-0254097P.  
XX 05-JAN-2001; 2001US-0259678P.  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI, 2001-488787/53.  
DR N-PSDB; AAS30231.  
XX  
PT New polynucleotides and polypeptides, useful for diagnosing, treating, preventing or prognosing e.g. kidney, cardiovascular, blood, electrolyte imbalance or neoplastic disorders, autoimmune diseases, cancers.  
XX  
XX Claim 1; SEQ ID NO 149; 506pp; English.  
PS  
XX  
XX The invention relates to novel nucleic acids and polypeptides useful for diagnosing, treating, preventing and/or prognosing disorders related to these polypeptides. The polynucleotides are especially useful in the diagnosis, prognosis, prevention and/or treatment of diseases which include kidney disorders (e.g. renal failure or nephritis), cardiovascular disorders (e.g. hypertension or myocardial infarction), blood disorders (e.g. anaemia or blood coagulation disorders), electrolyte imbalance disorders (e.g. hyponatraemia or hyperkalaemia), neoplastic disorders (e.g. nephroma or renal cell cancer), autoimmune diseases, cancers, inflammatory diseases, reproductive system disorders, endocrine disorders, neural activity and neurological disorders, wound

CC healing and respiratory disorders. AAU18644-AAU18715 represent the novel  
CC human renal and cardiovascular-associated amino acid sequences of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at: ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 76 AA;

Query Match 25.0%; Score 382; DB 4; Length 76;  
Best Local Similarity 95.9%; Pred. No. 8.2e-32;  
Matches 71; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 90 EYPQGLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVSYVQFHRSL 149  
Db :|||||  
1 KYPQGLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVSYVQFHRSL 60

QY 150 RTMSYRGTFQBFQR 163

Db 61 RTMSYRGTFQBFQR 74

RESULT 85

ABBI14904

ID ABBI14904 standard; protein; 76 AA.

AC ABBI14904;

XX

XX 23-JAN-2002 (first entry)

XX

DE Human nervous system related polypeptide SEQ ID NO 3561.

XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
KW antiparkinsonian; antisticking; antianaemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.

XX Homo sapiens.

XX WO200159063-A2.

XX

PD 16-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US001334.

XX

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236377P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 20-OCT-2000; 2000US-0242221P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.



```
PR 01-SEP-2000; 2000US-0223287P.
PR 01-SEP-2000; 2000US-0223343P.
PR 01-SEP-2000; 2000US-0223344P.
PR 01-SEP-2000; 2000US-0223345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229511P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 01-SEP-2000; 2000US-0223287P.
PR 01-SEP-2000; 2000US-0223343P.
PR 01-SEP-2000; 2000US-0223344P.
PR 01-SEP-2000; 2000US-0223345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229511P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249219P.
PR 17-NOV-2000; 2000US-0249220P.
PR 17-NOV-2000; 2000US-0249221P.
PR 17-NOV-2000; 2000US-0249222P.
PR 17-NOV-2000; 2000US-0249223P.
PR 17-NOV-2000; 2000US-0249224P.
PR 17-NOV-2000; 2000US-0249225P.
PR 17-NOV-2000; 2000US-0249226P.
PR 17-NOV-2000; 2000US-0249227P.
PR 17-NOV-2000; 2000US-0249228P.
PR 17-NOV-2000; 2000US-0249229P.
PR 17-NOV-2000; 2000US-0249230P.
PR 17-NOV-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251033P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251858P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254037P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764863.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-416807/39.
DR N-PSDB; ACD01486.
XX
XX New nucleic acids and polypeptides, useful for diagnosing, prognosing,
PT preventing and/or treating e.g. kidney, cardiovascular, blood,
PT electrolyte imbalance, neoplastic, respiratory, or neurological diseases
or disorders.
XX
PS Claim 11; Page 304; 363pp; English.
XX
XX The invention relates to isolated nucleic acids encoding novel
CC polypeptides. The nucleic acids are useful for chromosome mapping, for
CC radiation hybrid mapping, for detection of cancer, in gene therapy, for
CC identifying individuals from minute biological samples, as an alternative
CC to restriction fragment length polymorphism (RFLP) analysis, in forensic
CC biology and as hybridisation probes for differential identification of
CC tissues or cell types present in a biological sample. Compositions
CC comprising the polynucleotides, polypeptides and antibodies specific for
CC the polypeptides may be used in the diagnosis, prognosis, prevention
CC and/or treatment of kidney disorders (e.g. glomerulonephritis, urinary
CC tract infections, chronic nephritis), cardiovascular disorders (e.g.
CC arrhythmias, carotid heart disease, endocarditis), blood disorders
CC (e.g. thrombosis, anaemia, haemoglobin abnormalities), electrolyte
CC imbalance, neoplastic disorders (e.g. cancers), respiratory disorders
CC (e.g. acute rhinitis, sinusitis, pharyngitis) and neurological disorders
CC (e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease).
CC Sequences ABU97259-ABU97330 represent human polypeptides of the invention
XX
SQ Sequence 76 AA;
Query Match 25.0%; Score 382; DB 6; Length 76;
Best Local Similarity 95.9%; Pred. No. 8.2e-32;
Matches 71; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 90 EYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVWSYQFHRSL 149
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1 KYQPGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVWSYQFHRSL 60
```

Qy	150	RTWSYRGTFQFPCR 163             61 RTWSYRGXFOFCCR 74
Dd		
RESULT 87		
ABM84007		
ID	ABM84007 standard; protein; 254 AA.	
XX	AC	
XX	ABM84007;	
DT	18-NOV-2004 (first entry)	
DE	Human diagnostic and therapeutic pprotein SEQ ID NO:4256.	
KW	gene therapy; human diagnostic and therapeutic polynucleotide; dithp.	
OS	Homo sapiens.	
XX	WO2004023973-A2.	
Pn	25-MAR-2004.	
PD	12-SEP-2003; 2003WO-US028227.	
PF	12-SEP-2002; 2002US-0410259P.	
PR	12-SEP-2002; 2002US-0410260P.	
XX	(INCY-) INCYTE CORP.	
PA	Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F; Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV; Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP; Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH; Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL; Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES; Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D; Patury S, Shi X, Suarez CJ; WPI: 2004-329368/30. N-PSDB; ACN42659.	
DR	New diagnostic and therapeutic polynucleotides and polypeptides, useful in diagnosing a condition, disease or disorder associated with human molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or in gene mapping.	
XX	Claim 27; Page; 190pp; English.	
PS	The invention relates to novel diagnostic and therapeutic polynucleotides selected from one of the 2722 sequences defined in the specification. A polynucleotide of the invention may have a use in gene therapy. The human diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be used to diagnose a particular condition, disease or disorder associated with human molecules, e.g. cell proliferative disorders, autoimmune/inflammatory disorder, developmental disorder, endocrine disorder, neurological disorders, gastrointestinal disorders, or infections caused by virus, bacteria, fungi or parasite. The dithp molecules may also be used in genetic mapping, in identifying individuals from minute biological samples, in detecting single nucleotide polymorphisms, as molecular weight markers, and for somatic or germline gene therapy. The present sequence represents a dithp protein of the invention. Note: The sequence data for this patent is not represented in the printed specification, but was obtained in electronic format directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm	
XX	Sequence 254 AA;	
Qy	Query Match 24.6%; Score 376.5; DB 8; Length 254; Best Local Similarity 30.2%; Pred. No. 1.7e-30; Matches 83; Conservative 48; Mismatches 85; Indels 59; Gaps 4;	
OY	24 GVRLPPFCRGMKEIANFPVRPSDVIVTPKSGTSLLOEVVLVSOGADPDEIGLMVID 83	

CC molecules may also be used in genetic mapping, in identifying individuals  
 CC from minute biological samples, in detecting single nucleotide  
 CC polymorphisms, as molecular weight markers, and for somatic or germline  
 CC gene therapy. The present sequence represents a dithp protein of the  
 CC invention. Note: The sequence data for this patent is not represented in  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)

XX Sequence 254 AA;

Query Match 24.6%; Score 376.5; DB 8; Length 254;  
 Best Local Similarity 30.2%; Pred. No. 1.7e-30;  
 Matches 83; Conservative 48; Mismatches 85; Indels 59; Gaps 4;

QY 24 GVRPPFCRGKMEETANFVRPSDVWIVTPKSGTSLQEVVYLVSQADPDEIGLNMID 83  
 DB 17 GVPLIKYFAEALGPLQSFQARPDDLINTYPKSGTTWVSQILDIMYQGGDLEKCNRAPIY 76  
 QY 84 EQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYNARNPKDLVSY 143  
 DB 77 VRVPFLE-----VVYVARNPKDVAVSY 99

QY 144 QFHSLRTMSYRGTFQBFRCRFMNDKLGYSWFHVBQFWEHRMDSNVLFYEDMHRDL 203  
 DB 100 HFHRMEKAHPGPTWDSFLEKFMAGEVSYGSWYQHVQEWELSRTHPVLVLYFYEDMKENP 159

QY 204 VTMVEQLARFLGVSCKAQLAETHC-----HQLVDQCCNAEALP-V 245  
 DB 160 KREIQKILEFVGRSLPPEETDMFVQHTSFKEMKKNPMTNTYTPVQELMDHSIS-----PFM 215

QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQMKCDLTF 280  
 DB 216 RKGMDGDKTTFTVAQNERFDADYAEKMGCSLSF 250

RESULT 89

ABM84006  
 ID ABM84006 standard; protein; 254 AA.  
 XX  
 AC ABM84006;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE Human diagnostic and therapeutic pprotein SEQ ID NO:4255.  
 XX  
 KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004023973-A2.  
 XX  
 PD 25-MAR-2004.  
 XX  
 PF 12-SEP-2003; 2003WO-US028227.  
 XX  
 PR 12-SEP-2002; 2002US-0410259P.  
 PR 12-SEP-2002; 2002US-0410260P.  
 XX  
 PA (INCY-) INCYTE CORP.  
 XX  
 PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;  
 PI Harthshorne TA, Suchorolski WT, Altus CM, Pitts SJ, Elder LV;  
 PI Mooney EM, Deleane AM, Panesar IS, Banville SC, Reddy TP;  
 PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;  
 PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;  
 PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtan ES;  
 PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;  
 PI Patury S, Shi X, Suarez CJ;  
 XX  
 WIPI; 2004-329368/30.  
 DR N-PSDB; ACN42658.  
 XX  
 PT New diagnostic and therapeutic polynucleotides and polypeptides, useful

PT in diagnosing a condition, disease or disorder associated with human  
 PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or  
 PT in gene mapping.

XX Claim 27; Page: 190pp; English.

XX The invention relates to novel diagnostic and therapeutic polynucleotides  
 CC selected from one of the 2722 sequences defined in the specification. A  
 CC polynucleotide of the invention may have a use in gene therapy. The human  
 CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be  
 CC used to diagnose a particular condition, disease or disorder associated  
 CC with human molecules, e.g. cell proliferative disorders,  
 CC autoimmune/inflammatory disorder, developmental disorder, endocrine  
 CC disorder, neurological disorders, gastrointestinal disorders, or  
 CC infections caused by virus, bacteria, fungi or parasite. The dithp  
 CC molecules may also be used in genetic mapping, in identifying individuals  
 CC from minute biological samples, in detecting single nucleotide  
 CC polymorphisms, as molecular weight markers, and for somatic or germline  
 CC gene therapy. The present sequence represents a dithp protein of the  
 CC invention. Note: The sequence data for this patent is not represented in  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)

XX Sequence 254 AA;

Query Match 24.6%; Score 376.5; DB 8; Length 254;  
 Best Local Similarity 30.2%; Pred. No. 1.7e-30;  
 Matches 83; Conservative 48; Mismatches 85; Indels 59; Gaps 4;

QY 24 GVRPPFCRGKMEETANFVRPSDVWIVTPKSGTSLQEVVYLVSQADPDEIGLNMID 83  
 DB 17 GVPLIKYFAEALGPLQSFQARPDDLINTYPKSGTTWVSQILDIMYQGGDLEKCNRAPIY 76  
 QY 84 EQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYNARNPKDLVSY 143  
 DB 77 VRVPFLE-----VVYVARNPKDVAVSY 99

QY 144 QFHSLRTMSYRGTFQBFRCRFMNDKLGYSWFHVBQFWEHRMDSNVLFYEDMHRDL 203  
 DB 100 HFHRMEKAHPGPTWDSFLEKFMAGEVSYGSWYQHVQEWELSRTHPVLVLYFYEDMKENP 159

QY 204 VTMVEQLARFLGVSCKAQLAETHC-----HQLVDQCCNAEALP-V 245  
 DB 160 KREIQKILEFVGRSLPPEETDMFVQHTSFKEMKKNPMTNTYTPVQELMDHSIS-----PFM 215

QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQMKCDLTF 280  
 DB 216 RKGMDGDKTTFTVAQNERFDADYAEKMGCSLSF 250

RESULT 90

ABM82619  
 ID ABM82619 standard; protein; 254 AA.  
 XX  
 AC ABM82619;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE Human diagnostic and therapeutic pprotein SEQ ID NO:2868.  
 XX  
 KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004023973-A2.  
 XX  
 PD 25-MAR-2004.  
 XX  
 PF 12-SEP-2003; 2003WO-US028227.  
 XX  
 PR 12-SEP-2002; 2002US-0410259P.  
 PR 12-SEP-2002; 2002US-0410260P.  
 XX

```
PA (INCY-) INCYTE CORP.
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.
DR N-PSDB; ACN41271.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
XX Claim 27; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC polynucleotide and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
XX Sequence 254 AA;
XX
XX Query Match 24.6%; Score 376.5; DB 8; Length 254;
XX Best Local Similarity 30.2%; Pred. No. 1.7e-30;
XX Matches 83; Conservative 48; Mismatches 85; Indels 59; Gaps 4;
XX
Qy 24 GVRLPPFCRGKMEBIANFPVPSDVWIVTPKSGTSLLOEVVYLVSGADPDEIGLNMID 83
Db 17 GVPLIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDIMYQGGDLEKCNRAPIY 76
Qy 84 EQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDLVVSYY 143
Db 77 VRVPFLE-----VVYVARNPKDVAVSYY 99
Qy 144 QFHRSRLTMSYRGTFQBFRCRPMNDKLGYSWFEHVQEFWEHRMDSNVFLFLKYEDMHRDL 203
Db 100 HFHRMEKAHPPEPGTWDSEFLKFMAGEVSYGWSYQHVQEWELSRTHPVLVLYFYEDMKENP 159
Qy 204 VTWVEQLARFLGVSCDKAQLAEALTEHC-----HQLVDCCNAEALP-V 245
Db 160 KREIQKILEFVGRSLPEETMBDFMQHTSFEMKKNPMNTYITVQELMDHDSIS----PFM 215
Qy 246 GRGRVGLWKDIFTVSMNEKFLDLVTKQKMGKCDLTF 280
Db 216 RKGMAQGWKTTFTVAQNERFDADYAEKMGAGSLSF 250
XX
RESULT 91
ID ABM82617
XX ABM82617 standard; protein; 254 AA.
AC ABM82617;
XX
DT 18-NOV-2004 (first entry)
XX
```

```
DE Human diagnostic and therapeutic pprotein SEQ ID NO:2866.
XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX Homo sapiens.
XX WO2004023973-A2.
XX 25-MAR-2004.
XX
XX 12-SEP-2003; 2003WO-US028227.
XX
XX 12-SEP-2002; 2002US-0410259P.
XX 12-SEP-2002; 2002US-0410260P.
XX (INCY-) INCYTE CORP.
XX
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.
DR N-PSDB; ACN41269.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
XX Claim 27; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC polynucleotide and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
XX Sequence 254 AA;
XX
XX Query Match 24.6%; Score 376.5; DB 8; Length 254;
XX Best Local Similarity 30.2%; Pred. No. 1.7e-30;
XX Matches 83; Conservative 48; Mismatches 85; Indels 59; Gaps 4;
XX
Qy 24 GVRLPPFCRGKMEBIANFPVPSDVWIVTPKSGTSLLOEVVYLVSGADPDEIGLNMID 83
Db 17 GVPLIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDIMYQGGDLEKCNRAPIY 76
Qy 84 EQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDLVVSYY 143
Db 77 VRVPFLE-----VVYVARNPKDVAVSYY 99
Qy 144 QFHRSRLTMSYRGTFQBFRCRPMNDKLGYSWFEHVQEFWEHRMDSNVFLFLKYEDMHRDL 203
Db 100 HFHRMEKAHPPEPGTWDSEFLKFMAGEVSYGWSYQHVQEWELSRTHPVLVLYFYEDMKENP 159
Qy 204 VTWVEQLARFLGVSCDKAQLAEALTEHC-----HQLVDCCNAEALP-V 245
Db 100 HFHRMEKAHPPEPGTWDSEFLKFMAGEVSYGWSYQHVQEWELSRTHPVLVLYFYEDMKENP 159
Qy 204 VTWVEQLARFLGVSCDKAQLAEALTEHC-----HQLVDCCNAEALP-V 245
XX
```

```

Db      160 KREIQKILEFVGRSLPBEETMDFMVQHTSFKEMKKNPMTNTYTVPOELMDHSIS-----PFM 215
Qy      246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db      216 RKGWAGDKTFTTVAQNERFDADYAEKMGCSLSF 250

RESULT 92
ABM82618
ID      ABM82618 standard; protein; 254 AA.
AC      ABM82618;
XX
XX      18-NOV-2004 (first entry)
XX      Human diagnostic and therapeutic pprotein SEQ ID NO:2867.
XX      gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX      Homo sapiens.
XX      WO2004023973-A2.
XX      25-MAR-2004.
XX      12-SEP-2003; 2003WO-US028227.
XX      12-SEP-2002; 2002US-0410259P.
XX      12-SEP-2002; 2002US-0410260P.
XX      (INCY-) INCYTE CORP.
XX      Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
XX      Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
XX      Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
XX      Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
XX      Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
XX      Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtson ES;
XX      Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
XX      Patry S, Shi X, Suarez CJ;
XX      WPI; 2004-329368/30.
XX      N-PSDB; ACN41270.
XX      New diagnostic and therapeutic polynucleotides and polypeptides, useful
XX      in diagnosing a condition, disease or disorder associated with human
XX      molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX      in gene mapping.
XX      Claim 27; Page; 190pp; English.
XX      The invention relates to novel diagnostic and therapeutic polynucleotides
XX      selected from one of the 2722 sequences defined in the specification. A
XX      polynucleotide of the invention may have a use in gene therapy. The human
XX      diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
XX      used to diagnose a particular condition, disease or disorder associated
XX      with human molecules, e.g. cell proliferative disorders,
XX      autoimmune/inflammatory disorder, developmental disorder, endocrine
XX      disorder, neurological disorders, gastrointestinal disorders, or
XX      infections caused by virus, bacteria, fungi or parasite. The dithp
XX      molecules may also be used in genetic mapping, in identifying individuals
XX      from minute biological samples, in detecting single nucleotide
XX      polymorphisms, as molecular weight markers, and for somatic or germline
XX      gene therapy. The present sequence represents a dithp protein of the
XX      invention. Note: The sequence data for this patent is not represented in
XX      the printed specification, but was obtained in electronic format directly
XX      from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX      Sequence 254 AA;
XX
Query Match      24.6%; Score 376.5; DB 8; Length 254;
Best Local Similarity 30.2%; Pred. No. 1.7e-30;
Matches      83; Conservative 48; Mismatches 85; Indels 59; Gaps 4;

```

```

Qy      24 GVRLLPFCRCGRMBEIANFVRPSDVMIVTYPKSGTSLQLQBVVLVSQADPDEIGLMNID 83
Db      17 GVPLIKYFAEALGPLQSFQARPDLLINTYPKSGTTWSQLDMYIOGGDLKCNRAPIY 76
Qy      84 EQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDLVVSYY 143
Db      77 VRVPFLE-----VVVARNPKDVAVSYY 99
Qy      144 QFHRSLRTMSYRGTFQBFRRFMDNDKLGYSWFEHVFWEHHRMDSNVLFKLYEDMHRDL 203
Db      100 HFHRMEKAHPPEPGTWDSEFLKFMAGEVSYGSWYQHVQEWELSRTHFVLYLFYEDMKENP 159
Qy      204 VTWVEQLARFLGVSCDKAQLKALTEHC-----HQLVDQCCNAEALP-V 245
Db      160 KREIQKILEFVGRSLPBEETMDFMVQHTSFKEMKKNPMTNTYTVPOELMDHSIS-----PFM 215
Qy      246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db      216 RKGWAGDKTFTTVAQNERFDADYAEKMGCSLSF 250

RESULT 93
ABM82620
ID      ABM82620 standard; protein; 254 AA.
AC      ABM82620;
XX      18-NOV-2004 (first entry)
XX      Human diagnostic and therapeutic pprotein SEQ ID NO:2869.
XX      gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX      Homo sapiens.
XX      WO2004023973-A2.
XX      25-MAR-2004.
XX      12-SEP-2003; 2003WO-US028227.
XX      12-SEP-2002; 2002US-0410259P.
XX      12-SEP-2002; 2002US-0410260P.
XX      (INCY-) INCYTE CORP.
XX      Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
XX      Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
XX      Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
XX      Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
XX      Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
XX      Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtson ES;
XX      Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
XX      Patry S, Shi X, Suarez CJ;
XX      WPI; 2004-329368/30.
XX      N-PSDB; ACN41272.
XX      New diagnostic and therapeutic polynucleotides and polypeptides, useful
XX      in diagnosing a condition, disease or disorder associated with human
XX      molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX      in gene mapping.
XX      Claim 27; Page; 190pp; English.
XX      The invention relates to novel diagnostic and therapeutic polynucleotides
XX      selected from one of the 2722 sequences defined in the specification. A
XX      polynucleotide of the invention may have a use in gene therapy. The human
XX      diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
XX      used to diagnose a particular condition, disease or disorder associated
XX      with human molecules, e.g. cell proliferative disorders,
XX      autoimmune/inflammatory disorder, developmental disorder, endocrine
XX      disorder, neurological disorders, gastrointestinal disorders, or
XX      infections caused by virus, bacteria, fungi or parasite. The dithp
XX      molecules may also be used in genetic mapping, in identifying individuals
XX      from minute biological samples, in detecting single nucleotide
XX      polymorphisms, as molecular weight markers, and for somatic or germline
XX      gene therapy. The present sequence represents a dithp protein of the
XX      invention. Note: The sequence data for this patent is not represented in
XX      the printed specification, but was obtained in electronic format directly
XX      from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX      Sequence 254 AA;
XX
Query Match      24.6%; Score 376.5; DB 8; Length 254;
Best Local Similarity 30.2%; Pred. No. 1.7e-30;
Matches      83; Conservative 48; Mismatches 85; Indels 59; Gaps 4;

```



CC disorder, neurological disorders, gastrointestinal disorders, or  
CC infections caused by virus, bacteria, fungi or parasite. The dithp  
CC molecules may also be used in genetic mapping, in identifying individuals  
CC from minute biological samples, in detecting single nucleotide  
CC polymorphisms, as molecular weight markers, and for somatic or germline  
CC gene therapy. The present sequence represents a dithp protein of the  
CC invention. Note: The sequence data for this patent is not represented in  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)  
XX  
SQ Sequence 254 AA;

Query Match 24.6%; Score 376.5; DB 8; Length 254;  
Best Local Similarity 30.2%; Pred. No. 1.7e-30;  
Matches 83; Conservative 48; Mismatches 85; Indels 59; Gaps 4;

QY 24 GVRLPPFCRGKMEETANPPVPSDVWIVTPYKSGTSLLOEVVLYSQGADPDDEIGLMNID 83  
DB 17 GVPLIKYFAEALGPQSFQARPDLLINTYPSGTTWVSQILDMYQGGDLKCNRAPIY 76  
QY 84 EQLPVLEYPPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVWSY 143  
DB 77 VRVPFLE-----VYTVARNPKDVAVSYY 99  
QY 144 QHRSRLRTMSYRGTFQFCRRPMNDKLGYSWFHEHVFWEHRMDSNVLFKYEDMHRDL 203  
DB 100 HPHRMEKAHPPEGTWDSFLEKFMAGEVSYGSYQHVQEWELSRTHPVLVLYFYEDMKENP 159  
QY 204 VTMVQLARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAEALP-V 245  
DB 160 KREIQKILFVGRSLPEETMDPMFVQHTSPKEMKNPMTNTYTPQELMDHSIS----PPM 215  
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
DB 216 RKGMDGDKMTTFTVAQNERFDADYAEKMGCSLSF 250

RESULT 94  
ADG76311  
ID ADG76311 standard; protein; 259 AA.  
XX  
AC ADG76311;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Human Incyte enzymatic protein (SeqID 3).  
XX  
KW enzyme; human; biogenesis; biodegradation; autoimmune; inflammatory;  
KW infectious; neurological; cardiovascular; metabolic; cell proliferative;  
KW cancer; cytostatic; immunomodulator.  
XX  
OS Homo sapiens.  
XX  
PN WO2003083082-A2.  
XX  
PD 09-OCT-2003.  
XX  
PF 26-MAR-2003; 2003WO-US009772.  
XX  
PR 29-MAR-2002; 2002US-0368721P.  
PR 29-MAR-2002; 2002US-0368799P.  
XX  
PA (INCY-) INCYTE CORP.  
XX  
PI Tran UK, Kable AE, Ramkumar J, Ison CH, Richardson TW, Lee SY;  
PI Khare R, Marquis JP, Swarnakar A, Chawla NK, Elliott VS;  
PI Emerling BM, Becha SD, Hafalia AJA, Li JX, Griffin JA, Hawkins PR;  
PI Jin P, Chien D, Jiang X, Jackson AA, Mason PM, Bhatia UG;  
PI Burrill JD, Lee S, Blake JJ, Ho A, Zheng W, Gao J;  
XX  
DR WPI: 2003-804054/75.  
DR N-PSDB; ADG76366.  
XX

PT New human enzymes (ENZM) polypeptide, useful for preparing a composition  
PT for treating a disease associated with decreased expression or  
PT overexpression of ENZM e.g. cancer.  
XX  
PS Claim 58; SEQ ID NO 3; 385pp; English.  
XX  
CC This invention relates to novel isolated polynucleotides, and the enzymes  
CC encoded thereof, as well as any biologically active or immunogenic  
CC fragments. Specifically, it refers to human enzymes involved in the  
CC cellular processes of biogenesis and biodegradation, for example  
CC oxidoreductases, hydrolases and isomerases. The present invention  
CC describes the use of these nucleic acids and proteins (including single  
CC nucleotide polymorphisms) in the diagnosis, treatment and prevention of  
CC various disorders such as autoimmune/ inflammatory, infectious,  
CC neurological, cardiovascular, metabolic and cell proliferative such as  
CC cancer. Furthermore, the polypeptide is useful for preparing a  
CC composition for treating a disease associated with decreased expression  
CC or overexpression of the functional enzyme (ENZM) e.g. immune disorders.  
CC Accordingly, the compositions exhibit cytostatic and immunomodulator  
CC activities. This polypeptide sequence is a human incyte protein, an  
CC enzyme of the invention.  
XX  
SQ Sequence 259 AA;

Query Match 23.8%; Score 363.5; DB 7; Length 259;  
Best Local Similarity 29.6%; Pred. No. 4e-29;  
Matches 83; Conservative 53; Mismatches 91; Indels 53; Gaps 5;

QY 18 KYFEFHGVLRLPPFCRGKMEETANPPVPSDVWIVTPYKSGTSLLOEVVLYSQGADPDEI 77  
DB 12 KLKEVEGTLLOPATVDNWSQIQSFQAKPDLLICTYPRAGTTWIOEIVDMIEQNGDVEKC 71  
QY 78 GLMNIDEQLPVLEYPPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP 137  
DB 72 QRATIIQHRHPFIEWARP-----PQ-----PS-----ARNAKD 98  
QY 138 LVVSYQHFHRSRLRTMSYRGTFQFCRRPMNDKLGYSWFHEHVFWEHRMDSNVLFKYE 197  
DB 99 CMVSYHFFQRMNHLMPDEGTWEEYFETFKNGKVVWGSWFDHVKGWEMKDRHQILFLFY 158  
QY 198 DMHRDLVTMVQLARFLGVSCDKAQLEALTEH-----CHQLVDQCCNA 240  
DB 159 DIKDPKHEIRKVMQFMGKKVDEIVLDKIVQETSEKMKENPMTNRTSVTSKILDSISS 218  
QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
DB 219 F---MRKGTVGDKNKHFTVAQNERFDEIYRRKMGCTSF 255

RESULT 95  
ABB61517  
ID ABB61517 standard; protein; 338 AA.  
XX  
AC ABB61517;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 11343.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX

PA (PEKE ) PE CORP NY.  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI; 2001-656860/75.  
 XX N-PSDB; ABL05620.  
 DR  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 XX  
 PS Disclosure; SEQ ID NO 11343; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 338 AA;  
 Query Match 23.6%; Score 360.5; DB 4; Length 338;  
 Best Local Similarity 31.0%; Pred. No. 1.2e-28;  
 Matches 89; Conservative 51; Mismatches 112; Indels 35; Gaps 7;  
 QY 9 PSTP--GFESKYFFHGVRLPPFCCKGMEIATFPVPSPDVWIVTYPKSGTSLLOEVVY 66  
 Db 34 PSVPVVGWNEQRF-----CRLADTQPVLDRLVDFEVRDDWIVTLKCGTITWQELAW 88  
 QY 67 LVSQGADPDEIGLMNIDBOLPVLEY-----PQGLDIIKELTSPRLIKSHLPYRFLPS 119  
 Db 89 LVINECDFETAKSVDLTHRSPLFEGVVPVPHDTIAAANALPSRLIKSHLPAMWLP 148  
 QY 120 DLHNGDSKVIYMARPKDLVSYQYFHSRLTMSYRGTFQFCRPFPMNDKLGYSWFHV 179  
 Db 149 QIWSKRPKIIVYRNPKDAAISYFHHWGM--VGQGTKSDPMHGFIDGYVNFPCWPHI 206  
 QY 180 QBFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCKAQLBALTEHC--HOLVQD- 236  
 Db 207 LDFWQLRHREPNIFFTSYERMKQLCQVISEVAQFLERSVSQEQMQMQORHLSFESMRNP 266  
 QY 237 CCN-----AEALPVGRGRVGLWKDIFTVSMNEKFDL 267  
 Db 267 ACNHVKEFESMKAAGREVEEPFRVRRGVVGSHKDELTAIIRFIDL 313

RESULT 96  
 ABB62612  
 ID ABB62612 standard; protein; 346 AA.  
 AC ABB62612;  
 XX  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 14628.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.  
 XX (PEKE ) PE CORP NY.  
 PA Venter JC, Adams M, Li PWD, Myers EW;  
 XX WPI; 2001-656860/75.  
 DR N-PSDB; ABL06715.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 XX  
 PS Disclosure; SEQ ID NO 14628; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 346 AA;  
 Query Match 23.5%; Score 359.5; DB 4; Length 346;  
 Best Local Similarity 32.3%; Pred. No. 1.6e-28;  
 Matches 82; Conservative 35; Mismatches 88; Indels 49; Gaps 5;  
 QY 22 FHGVR-----LPPFCRGKMEIATFPVPSPDVWIVTYPKSGTSLLOEVVILVS 69  
 Db 27 FHGERTGVQVSGYGFPHKYKDEAERYNFARPDWVIATVPRSGTTWTQELIWLVA 86  
 QY 70 QGADPDEIGLMNIDBOLPVLEY-----PQGLDIIK 100  
 Db 87 NGLDFEHAQERPLTERPFFPFFPLFVHPKIEELQENRDSAEALEFTEKTARFGYEALS 146  
 QY 101 EL--TSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDLVSYQYFHSRLTMSYRGTF 158  
 Db 147 EIPRSQRRFIKTHFPFSLMPPSVLEKKCKVIYVVRDPKDVAVSYHNLRLFRFTQGYVGDF 206  
 QY 159 QFCRCERNMDKLGYSW---FEHVQEFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLG 215  
 Db 207 ERYWHYFQN---GLNPMLPYYSHVKEAREHAHLNSVLFRLRYEDMLADLPGAINSTASFL 263  
 QY 216 VSCDKAQLALTEH 229  
 Db 264 CPPKPEDMDRLLDH 277  
 RESULT 97  
 ABB82695  
 ID ABB82695 standard; protein; 254 AA.  
 XX  
 XX  
 AC ABB82695;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE Human diagnostic and therapeutic pprotein SEQ ID NO:2944.  
 XX  
 KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004023973-A2.  
 XX  
 PD 25-MAR-2004.  
 XX  
 PF 12-SEP-2003; 2003WO-US028227.  
 XX

```
PR 12-SEP-2002; 2002US-0410259P.
XX 12-SEP-2002; 2002US-0410260P.
XX (INCY-) INCYTE CORP.
XX
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patry S, Shi X, Suarez CJ;
XX
XX WPI; 2004-329368/30.
DR N-PSDB; ACN41347.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
XX Claim 27; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC polynucleotide and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorders, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
XX Sequence 254 AA;
SQ
Query Match 23.4%; Score 357.5; DB 8; Length 254;
Best Local Similarity 29.9%; Pred. No. 1.7e-28;
Matches 81; Conservative 46; Mismatches 93; Indels 51; Gaps 4;
QY 24 GVRLPFFCRGMEETANFVRPSDVWIVTYPKSGTSLQEVVLYSQGADPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSFQARPDDLITSTYPKSGTTWVSQILDMIYQGGD----- 66
QY 84 EQLPVLEYPPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVYIYARNPKDLVWSY 143
DB 67 -----LEKCHR-----VPQTLDDQKVVKVYVARNAKDVAVSYY 99
QY 144 QFHRSLRTMSYRGTFQEFRCFRFMDNDKLGYSWFEHVFQEFWEHRMDSNVFLFKYEDMHRDL 203
DB 100 HFYHMAKVHPPEFTWDSFLEKFMVGEVSGYSWYQHVQEWELSRTHPVLVLYFYEDMKENP 159
QY 204 VTMVQLARFLGVSCDKAQLAETHC---HQLVQCCNAEALP-----VGRGR 249
DB 160 KREIQKILEFVGRSLPEETVDFVVOHTSPFKEMKKNPMTNTYTVQEFMDHSISPFMRKGM 219
QY 250 VGLWKDIFTVSNKEKFDLVYKMKGKCDLTF 280
DB 220 AGDKTFTTVAQNERFDADYAEKMACCSLSF 250
RESULT 98
ABM82697
ID ABM82697 standard; protein; 254 AA.
XX
AC ABM82697;
```

```
XX 18-NOV-2004 (first entry)
XX
XX Human diagnostic and therapeutic pprotein SEQ ID NO:2946.
XX
XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
XX Homo sapiens.
XX
XX WO2004023973-A2.
XX
XX 25-MAR-2004.
XX
XX 12-SEP-2003; 2003WO-US028227.
XX
XX 12-SEP-2002; 2002US-0410259P.
XX 12-SEP-2002; 2002US-0410260P.
XX (INCY-) INCYTE CORP.
XX
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patry S, Shi X, Suarez CJ;
XX
XX WPI; 2004-329368/30.
DR N-PSDB; ACN41349.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
XX Claim 27; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC polynucleotide and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorders, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
XX Sequence 254 AA;
SQ
Query Match 23.4%; Score 357.5; DB 8; Length 254;
Best Local Similarity 29.9%; Pred. No. 1.7e-28;
Matches 81; Conservative 46; Mismatches 93; Indels 51; Gaps 4;
QY 24 GVRLPFFCRGMEETANFVRPSDVWIVTYPKSGTSLQEVVLYSQGADPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSFQARPDDLITSTYPKSGTTWVSQILDMIYQGGD----- 66
QY 84 EQLPVLEYPPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVYIYARNPKDLVWSY 143
DB 67 -----LEKCHR-----VPQTLDDQKVVKVYVARNAKDVAVSYY 99
QY 144 QFHRSLRTMSYRGTFQEFRCFRFMDNDKLGYSWFEHVFQEFWEHRMDSNVFLFKYEDMHRDL 203
DB 100 HFYHMAKVHPPEFTWDSFLEKFMVGEVSGYSWYQHVQEWELSRTHPVLVLYFYEDMKENP 159
```

QY 204 VTMVEQLARFLGVSCDKAQLAETHC---HQLVDQCCNAALP-----VGRGR 249  
 Db 160 KREIQKILEFVGRSLFEETVDFVVOHTSFKEKMKPNMTNTTVPOEFMDHSISPMPKGM 219  
 QY 250 VGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
 Db 220 AGDWKTTTVAQNERFDADYAEKMGCSLSF 250

RESULT 99  
 ID ADD48500 standard; protein; 283 AA.  
 XX  
 AC ADD48500;  
 XX  
 DT 02-DEC-2004 (revised)  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Rat Protein CAA45007, SEQ ID NO 14201.  
 XX  
 XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
 XX  
 OS Rattus norvegicus.  
 OS Unidentified.  
 XX  
 PN WO2003016475-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 XX 14-AUG-2002; 2002WO-US025765.  
 XX  
 PR 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX  
 XX (GEO) GEN HOSPITAL CORP.  
 PA (FAR) BAYER AG.  
 XX  
 XX Woolf C, D'urzo D, Befort K, Costigan M;  
 XX WPI; 2003-268312/26.  
 DR GENBANK; CAA45007.  
 XX  
 XX New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX  
 XX Example 1; Page; 1017pp; English.  
 PS  
 XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (described in Table 3

CC of the specification) which is differentially expressed during pain.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 283 AA;

Query Match 23.4%; Score 357.5; DB 7; Length 283;  
 Best Local Similarity 33.5%; Pred. No. 1.9e-28;  
 Matches 92; Conservative 53; Mismatches 105; Indels 25; Gaps 8;

QY 19 YFEFHGVRLPDFCRGK--MERIAN-FPVRPDSVMIVTPKSGTSLLOEVVYLVSGADPD 75  
 Db 3 YTFEGIPFPFAFGIPKGTQLQVNCNKFVVKBEDLLLTYPKSGTNWLIIEIVCLIQTKGDPK 62  
 QY 76 EIGLWNIDEQLPVILEYPOPLDIKELTSPRLIKSHLPYRFLPSDLHNGDSKVYMARNP 135  
 Db 63 WIQSVTIWDRSPWIE-TDLGYDMLIKKKGPRLITSHLPMLFSLFKSKAKVIYLIRNP 121  
 QY 136 KDLVVVYVQFHRSLRTMSYR---GTFQFCRRFMNDKLGYSWFHVEHVOEFMEHRMDSNV 191  
 Db 122 RDLVSGYFYF--GKTTLAKPDSLGTIVEM---FLKGVYPYGSWFHRAWLSMRELDNF 177  
 QY 192 LFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLAETHCHQLVDQCCNAE-----A 242  
 Db 178 LLLYEDMKDVTMTGTTIKKICDFLCKLEPDELVLVYSSFQVMKENNMYNLMWEKELI 237  
 QY 243 LP---VGRGRVGLWKDIFTVSMNEKFDLVYKQKM 273  
 Db 238 LPGAFTMRNGTTGDWKNHFTVAQAEAFDKVFOEKM 272

RESULT 100

AEF05774  
 ID AEF05774 standard; protein; 285 AA.

XX AC AEF05774;

XX DT 23-FEB-2006 (first entry)

XX DE Human hydroxysteroid sulfotransferase.

XX KW drug screening; metabolism; DNA typing;

XX KW hydroxysteroid sulfotransferase 2A1; SULT2A1; enzyme.

XX OS Homo sapiens.

XX PN WO2005123922-A1.

XX PD 29-DEC-2005.

XX PF 17-JUN-2005; 2005WO-CA000941.

XX PR 17-JUN-2004; 2004US-0580540P.

XX PA (UYGU-) UNIV GUELPH.

XX PI Squires J, Sinclair P;

XX WPI; 2006-067510/07.

XX New polynucleotides, useful in screening for a substance that enhances  
 PT metabolism of 16-androstene steroids or in reducing or preventing boar  
 PT taint in a pig.

XX PS Disclosure; Fig 5; 93pp; English.

XX The new invention relates to a polynucleotide encoding a  
 CC sulfotransferase. The sequence comprises a fully defined 1036-bp sequence  
 CC (SEQ ID NO: 1); sequence encoding a protein having a fully defined 286-  
 CC amino acid sequence (SEQ ID NO: 2); a sequence encoding a  
 CC sulfotransferase that hybridizes to the sequence in (A); or a sequence  
 CC having at least about 90% identity with the sequences in (A). Also

CC described are a vector comprising the polynucleotide operatively linked  
CC to a regulatory sequence; a recombinant host cell comprising the vector;  
CC a polypeptide; an antibody that specifically binds to the polypeptide; a  
CC method of determining the genetic capacity of a male pig to develop boar  
CC taint; a method for reducing or preventing boar taint in a pig; a method  
CC for screening for a substance that enhances metabolism of 16-androstene  
CC steroids in a pig; a method for screening for a substance that enhance  
CC metabolism of 16-androstene steroids by enhancing transcription and/or  
CC translation of the nucleotide sequence encoding sulfotransferase 2A1  
CC (SULT2A1); and a method of genetically typing pigs to determine those  
CC with desired boar taint characteristics. The isolated polynucleotide  
CC comprises a nucleotide change of a cytosine (C) to a thymine (T) at  
CC nucleotide position 219. The polypeptide is a SULT2A1 hydroxysteroid  
CC sulfotransferase. The polynucleotide is useful in screening for a  
CC substance that enhances metabolism of 16-androstene steroids or in  
CC reducing or preventing boar taint in a pig. The present sequence is a  
CC human hydroxysteroid sulfotransferase variant.  
XX

SQ	Sequence 285 AA;	
Query Match	23.3%;	Score 356.5; DB 10; Length 285;
Best Local Similarity	31.0%;	Pred. No. 2.5e-28;
Matches	83; Conservative 59; Mismatches 109; Indels 17; Gaps 5;	
QY	22 FHGVRLLP--FCRGRMEEIAN--PPVRPSDVWIVTPKSGTSLQEVVYLVSQADPDDEIG	78
Db	8 FEGIAFPTMGFRSETLRKVRDEVFIRDEVDVILITPKSGTNWLABILCLMHSGDKAKWITQ	67
QY	79 LMNIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKD	138
Db	68 SVPIWERSPWVE-SGIGYTALSETSPRLFSSHLPIQLFPKSFPSKAKVIYLMENPRDV	126
QY	139 VVSYYQFHRSLRTMSYRGTFQBFCCRFMNDKLGYSWFHVGQEFWEHRMDSNVLFKYED	198
Db	127 LVSGYFFWKNMKFIKKPKSWSEYFEWFQGTVVYGSWFEDIHGWMPMREKNFLLSYEE	186
QY	199 MHRDLVTWVEQLARFLGVSCDKAQLALETHC---HQLVDQCCNAEALPVG-----	246
Db	187 LKQDTGRTIEKICQFLGKTLPEPEELNLILKNSSFQSMKENKMSNYSLLSVDYVVDKAQLL	246
QY	247 -RGRVGLWKDIFTVSMNEKFDLVYKOKM	273
Db	247 RKGVSQDNKNHFTVAQAEFDKLFQEKM	274

Search completed: May 17, 2006, 10:59:05  
Job time : 211 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2006, 11:05:51 ; Search time 601 Seconds  
(without alignments)  
721.557 Million cell updates/sec

Title: US-10-768-158-2

Perfect score: 1530

Sequence: 1 MAESEAETPTGPEFESKYP.....FDLVYKQKMGKCDLTFDFYL 284

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 8366291 seqs, 1526956180 residues

Total number of hits satisfying chosen parameters: 8366291

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Pending Patents\_AA\_Main.\*

1: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US01-41986-4 COMB.pap.\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US02-11654-60 COMB.pap.\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US02-11654-81 COMB.pap.\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US02-11654-103 COMB.pap.\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US02-11654-104 COMB.pap.\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
9: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
10: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
11: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
12: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
13: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
14: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
15: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
16: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
17: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
18: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
19: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
20: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
21: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
22: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
23: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
24: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
25: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
26: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
27: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
28: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
29: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
30: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
31: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
32: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
33: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
34: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
35: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
36: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
37: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
38: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
39: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
40: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
41: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
42: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*  
43: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US04-02851-2 COMB.pap.\*

44: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US114\_COMB.pap.\*  
45: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US600\_COMB.pap.\*  
46: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US601\_COMB.pap.\*  
47: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US602\_COMB.pap.\*  
48: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US603\_COMB.pap.\*  
49: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US604\_COMB.pap.\*  
50: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US605\_COMB.pap.\*  
51: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US606\_COMB.pap.\*  
52: /EMC\_Celerra\_SIDS3/ptodata/2/paa/US607\_COMB.pap.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1530	100.0	284	1	PCT-US01-41986-4	Sequence 4, Appli
2	1530	100.0	284	1	PCT-US01-41986-6	Sequence 6, Appli
3	1530	100.0	284	1	PCT-US02-11654-60	Sequence 60, Appli
4	1530	100.0	284	1	PCT-US02-11654-81	Sequence 81, Appli
5	1530	100.0	284	1	PCT-US02-11654-103	Sequence 103, Appli
6	1530	100.0	284	1	PCT-US02-11654-104	Sequence 104, Appli
7	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
8	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
9	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
10	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
11	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
12	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
13	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
14	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
15	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
16	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
17	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
18	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
19	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
20	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
21	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
22	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
23	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
24	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
25	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
26	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
27	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
28	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
29	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
30	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
31	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
32	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
33	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
34	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
35	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
36	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
37	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
38	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
39	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
40	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
41	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
42	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
43	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
44	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
45	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
46	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
47	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
48	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
49	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
50	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
51	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
52	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
53	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli
54	1530	100.0	284	1	PCT-US04-02851-2	Sequence 2, Appli





```
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; FILE REFERENCE: 51158-20040.40
; CURRENT APPLICATION NUMBER: PCT/US02/11654
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-11654-60

Query Match      100.0%; Score 1530; DB 1; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEATPTPGEFESKYFFHGVRLPPFCRGKMEIEIANFVRPSDVWIVTPKSGTSL 60
Db      1 MAESEATPTPGEFESKYFFHGVRLPPFCRGKMEIEIANFVRPSDVWIVTPKSGTSL 60

Qy      61 LQEVVYLVSQADPDEIGLNMIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LQEVVYLVSQADPDEIGLNMIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYMARNP KD LVSYQPHRS LR TMS YRG TQEF CRR FMND KLG YGS WFEHVQ 180
Db      121 LHNGDSKVIYMARNP KD LVSYQPHRS LR TMS YRG TQEF CRR FMND KLG YGS WFEHVQ 180

Qy      181 EFWHRMDSNVLF LK YED MHRDLVTWVEQLARFLGVSCDKAQL EALTEHCHQLVDQCCNA 240
Db      181 EFWHRMDSNVLF LK YED MHRDLVTWVEQLARFLGVSCDKAQL EALTEHCHQLVDQCCNA 240

Qy      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTDFYL 284
Db      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTDFYL 284
```

```
RESULT 4
PCT-US02-11654-81
; Sequence 81, Application PC/TUS0211654
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; FILE REFERENCE: 51158-20040.40
; CURRENT APPLICATION NUMBER: PCT/US02/11654
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-11654-103

Query Match      100.0%; Score 1530; DB 1; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEATPTPGEFESKYFFHGVRLPPFCRGKMEIEIANFVRPSDVWIVTPKSGTSL 60
```

```
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-11654-81

Query Match      100.0%; Score 1530; DB 1; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEATPTPGEFESKYFFHGVRLPPFCRGKMEIEIANFVRPSDVWIVTPKSGTSL 60
Db      1 MAESEATPTPGEFESKYFFHGVRLPPFCRGKMEIEIANFVRPSDVWIVTPKSGTSL 60

Qy      61 LQEVVYLVSQADPDEIGLNMIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LQEVVYLVSQADPDEIGLNMIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYMARNP KD LVSYQPHRS LR TMS YRG TQEF CRR FMND KLG YGS WFEHVQ 180
Db      121 LHNGDSKVIYMARNP KD LVSYQPHRS LR TMS YRG TQEF CRR FMND KLG YGS WFEHVQ 180

Qy      181 EFWHRMDSNVLF LK YED MHRDLVTWVEQLARFLGVSCDKAQL EALTEHCHQLVDQCCNA 240
Db      181 EFWHRMDSNVLF LK YED MHRDLVTWVEQLARFLGVSCDKAQL EALTEHCHQLVDQCCNA 240

Qy      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTDFYL 284
Db      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTDFYL 284
```

```
RESULT 5
PCT-US02-11654-103
; Sequence 103, Application PC/TUS0211654
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; FILE REFERENCE: 51158-20040.40
; CURRENT APPLICATION NUMBER: PCT/US02/11654
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-11654-103

Query Match      100.0%; Score 1530; DB 1; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEATPTPGEFESKYFFHGVRLPPFCRGKMEIEIANFVRPSDVWIVTPKSGTSL 60
```

```
Db 1 MAESEAETPTGPFESKYFHFHGVRLPPFCRGKMEIEANFPVRPSDWMIVTPKSGTSL 60
Qy 61 LOEVVYLVSQADPDEIGLNMIDQLVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLNMIDQLVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYMARNP KDVLVSYQPHRSRTMSYRGTFQFCRRPMNDKLGYSWFEHVQ 180
Db 121 LHNGDSKVIYMARNP KDVLVSYQPHRSRTMSYRGTFQFCRRPMNDKLGYSWFEHVQ 180
Qy 181 EFWHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 240
Db 181 EFWHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284

RESULT 6
PCT-US02-11654-104
; Sequence 104, Application PC/TUS0211654
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; TITLE OF INVENTION: in the Detection and Treatment of Various Cancers
; FILE REFERENCE: 51158-20040.40
; CURRENT APPLICATION NUMBER: PCT/US02/11654
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-11654-104

Query Match 100.0%; Score 1530; DB 1; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEAETPTGPFESKYFHFHGVRLPPFCRGKMEIEANFPVRPSDWMIVTPKSGTSL 60
Db 1 MAESEAETPTGPFESKYFHFHGVRLPPFCRGKMEIEANFPVRPSDWMIVTPKSGTSL 60
Qy 61 LOEVVYLVSQADPDEIGLNMIDQLVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLNMIDQLVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYMARNP KDVLVSYQPHRSRTMSYRGTFQFCRRPMNDKLGYSWFEHVQ 180
Db 121 LHNGDSKVIYMARNP KDVLVSYQPHRSRTMSYRGTFQFCRRPMNDKLGYSWFEHVQ 180
Qy 181 EFWHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 240
Db 181 EFWHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284
```

```
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284

RESULT 7
PCT-US02-11654-708
; Sequence 708, Application PC/TUS0211654
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; TITLE OF INVENTION: in the Detection and Treatment of Various Cancers
; FILE REFERENCE: 51158-20040.40
; CURRENT APPLICATION NUMBER: PCT/US02/11654
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 708
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-11654-708

Query Match 100.0%; Score 1530; DB 1; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEAETPTGPFESKYFHFHGVRLPPFCRGKMEIEANFPVRPSDWMIVTPKSGTSL 60
Db 1 MAESEAETPTGPFESKYFHFHGVRLPPFCRGKMEIEANFPVRPSDWMIVTPKSGTSL 60
Qy 61 LOEVVYLVSQADPDEIGLNMIDQLVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLNMIDQLVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYMARNP KDVLVSYQPHRSRTMSYRGTFQFCRRPMNDKLGYSWFEHVQ 180
Db 121 LHNGDSKVIYMARNP KDVLVSYQPHRSRTMSYRGTFQFCRRPMNDKLGYSWFEHVQ 180
Qy 181 EFWHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 240
Db 181 EFWHRMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284

RESULT 8
PCT-US04-00750-110
; Sequence 110, Application PC/TUS0400750
; GENERAL INFORMATION:
; APPLICANT: Karicheti, Venkateswarlu
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Eliassof, Scott D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,
; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
; TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 44373, 51164,
; TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
```



```
Db 1 MAESEATPTPGESKPYFHHGVRLLPPFCRGKMEIEANFVRPSDVWIVTPKSGTSL 60
Qy 61 LOEVVYLVSQADPDEIGLWNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLWNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVVQ 180
Db 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVVQ 180
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284

RESULT 11
US-10-121-024B-60
; Sequence 60, Application US/10121024B
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024B
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-024B-60

Query Match 100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEATPTPGESKPYFHHGVRLLPPFCRGKMEIEANFVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPTPGESKPYFHHGVRLLPPFCRGKMEIEANFVRPSDVWIVTPKSGTSL 60
Qy 61 LOEVVYLVSQADPDEIGLWNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLWNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVVQ 180
Db 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVVQ 180
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284

RESULT 12
US-10-121-024B-81
; Sequence 81, Application US/10121024B
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024B
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-024B-81

Query Match 100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEATPTPGESKPYFHHGVRLLPPFCRGKMEIEANFVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPTPGESKPYFHHGVRLLPPFCRGKMEIEANFVRPSDVWIVTPKSGTSL 60
Qy 61 LOEVVYLVSQADPDEIGLWNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLWNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVVQ 180
Db 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVVQ 180
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284

RESULT 13
US-10-121-024B-103
; Sequence 103, Application US/10121024B
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
```

```
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284

RESULT 12
US-10-121-024B-81
; Sequence 81, Application US/10121024B
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024B
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-024B-81

Query Match 100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEATPTPGESKPYFHHGVRLLPPFCRGKMEIEANFVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPTPGESKPYFHHGVRLLPPFCRGKMEIEANFVRPSDVWIVTPKSGTSL 60
Qy 61 LOEVVYLVSQADPDEIGLWNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLWNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVVQ 180
Db 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVVQ 180
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284

RESULT 13
US-10-121-024B-103
; Sequence 103, Application US/10121024B
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
```

```
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Useful in the Detection and Treatment of Various Cancers
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024B
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-024B-103

Query Match      100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEETIANPVRPSDVWIVTPKSGTSL 60
Db      1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEETIANPVRPSDVWIVTPKSGTSL 60

Qy      61 LOEVVYLVSQGADPEIGLNMNIDEQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LOEVVYLVSQGADPEIGLNMNIDEQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180
Db      121 LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180

Qy      181 EFWHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Db      181 EFWHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

Qy      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284

RESULT 14
US-10-121-024B-104
; Sequence 104, Application US/10121024B
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Useful in the Detection and Treatment of Various Cancers
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024B
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 284
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-10-121-024B-104

Query Match      100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEETIANPVRPSDVWIVTPKSGTSL 60
Db      1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEETIANPVRPSDVWIVTPKSGTSL 60

Qy      61 LOEVVYLVSQGADPEIGLNMNIDEQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LOEVVYLVSQGADPEIGLNMNIDEQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180
Db      121 LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180

Qy      181 EFWHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Db      181 EFWHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

Qy      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284

RESULT 15
US-10-121-024B-708
; Sequence 708, Application US/10121024B
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Useful in the Detection and Treatment of Various Cancers
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024B
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 708
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-024B-708

Query Match      100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEETIANPVRPSDVWIVTPKSGTSL 60
Db      1 MAESEATPSTPGFESKYFEFHGVRLLPPFCRGKMEETIANPVRPSDVWIVTPKSGTSL 60

Qy      61 LOEVVYLVSQGADPEIGLNMNIDEQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LOEVVYLVSQGADPEIGLNMNIDEQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180
Db      121 LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVQ 180
```

```
Db 121 LHNGDSKVIYARNPKDLVSYQFHRSLRTWTSYAGTQFECRRPMNDKLGSGWFEHVQ 180
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284
```

## RESULT 16

```
US-10-121-024C-60
; Sequence 60, Application US/10121024C
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; TITLE OF INVENTION: in the Detection and Treatment of Various Cancers
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024C
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-024C-60
```

```
Query Match 100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEATPSTPGEFESKYFEGHGVRLPPFCRGKMEELANPPVRPSDVWIVTYPKSGTSL 60
Db 1 MAESEATPSTPGEFESKYFEGHGVRLPPFCRGKMEELANPPVRPSDVWIVTYPKSGTSL 60

Qy 61 LQEVVYLVSQGADPDEIGLWNIDEQLPVLEYQPQGLDIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLVSQGADPDEIGLWNIDEQLPVLEYQPQGLDIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYARNPKDLVSYQFHRSLRTWTSYAGTQFECRRPMNDKLGSGWFEHVQ 180
Db 121 LHNGDSKVIYARNPKDLVSYQFHRSLRTWTSYAGTQFECRRPMNDKLGSGWFEHVQ 180

Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284
```

## RESULT 17

```
US-10-121-024C-81
; Sequence 81, Application US/10121024C
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
```

```
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; TITLE OF INVENTION: in the Detection and Treatment of Various Cancers
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024C
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-024C-81
```

```
Query Match 100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEATPSTPGEFESKYFEGHGVRLPPFCRGKMEELANPPVRPSDVWIVTYPKSGTSL 60
Db 1 MAESEATPSTPGEFESKYFEGHGVRLPPFCRGKMEELANPPVRPSDVWIVTYPKSGTSL 60

Qy 61 LQEVVYLVSQGADPDEIGLWNIDEQLPVLEYQPQGLDIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLVSQGADPDEIGLWNIDEQLPVLEYQPQGLDIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYARNPKDLVSYQFHRSLRTWTSYAGTQFECRRPMNDKLGSGWFEHVQ 180
Db 121 LHNGDSKVIYARNPKDLVSYQFHRSLRTWTSYAGTQFECRRPMNDKLGSGWFEHVQ 180

Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTTFDFYL 284
```

## RESULT 18

```
US-10-121-024C-103
; Sequence 103, Application US/10121024C
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; TITLE OF INVENTION: in the Detection and Treatment of Various Cancers
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024C
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
```

```
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-024C-103

Query Match      100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESAETPTSPGEPESKYFFHGVRLPPFCRGKMEIANTFVRPSDVIWVTPKSGTSL 60
Db 1 MAESAETPTSPGEPESKYFFHGVRLPPFCRGKMEIANTFVRPSDVIWVTPKSGTSL 60

Qy 61 LOEVVYLVSQGADPDEIGLNMNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQGADPDEIGLNMNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQ 180
Db 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQ 180

Qy 181 EFWEHRMDSNVLF LKYEDEMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLF LKYEDEMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFDFYL 284

RESULT 19
US-10-121-024C-104
; Sequence 104, Application US/10121024C
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Robert B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024C
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-024C-104

Query Match      100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESAETPTSPGEPESKYFFHGVRLPPFCRGKMEIANTFVRPSDVIWVTPKSGTSL 60
Db 1 MAESAETPTSPGEPESKYFFHGVRLPPFCRGKMEIANTFVRPSDVIWVTPKSGTSL 60

Qy 61 LOEVVYLVSQGADPDEIGLNMNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQGADPDEIGLNMNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQ 180
Db 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQ 180

Qy 181 EFWEHRMDSNVLF LKYEDEMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLF LKYEDEMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFDFYL 284

RESULT 19
US-10-121-024C-104
; Sequence 104, Application US/10121024C
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Robert B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024C
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-024C-104

Query Match      100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESAETPTSPGEPESKYFFHGVRLPPFCRGKMEIANTFVRPSDVIWVTPKSGTSL 60
Db 1 MAESAETPTSPGEPESKYFFHGVRLPPFCRGKMEIANTFVRPSDVIWVTPKSGTSL 60

Qy 61 LOEVVYLVSQGADPDEIGLNMNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQGADPDEIGLNMNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQ 180
Db 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQ 180

Qy 181 EFWEHRMDSNVLF LKYEDEMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLF LKYEDEMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFDFYL 284

RESULT 20
US-10-121-024C-708
; Sequence 708, Application US/10121024C
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Robert B.
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful
; FILE REFERENCE: 51158-20040.00
; CURRENT APPLICATION NUMBER: US/10/121,024C
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 708
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-024C-708

Query Match      100.0%; Score 1530; DB 31; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESAETPTSPGEPESKYFFHGVRLPPFCRGKMEIANTFVRPSDVIWVTPKSGTSL 60
Db 1 MAESAETPTSPGEPESKYFFHGVRLPPFCRGKMEIANTFVRPSDVIWVTPKSGTSL 60

Qy 61 LOEVVYLVSQGADPDEIGLNMNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQGADPDEIGLNMNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQ 180
Db 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFHVVQ 180

Qy 181 EFWEHRMDSNVLF LKYEDEMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLF LKYEDEMHRDLVTWVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFDFYL 284
```

## RESULT 21

US-10-427-631-7  
; Sequence 7, Application US/10427631  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE CORPORATION; TANG, Y. TOM;  
; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;  
; APPLICANT: BAUGHN, Mariah R.; LAL, Preeti G.;  
; APPLICANT: YUE, Henry; HILLMAN, Jennifer L.;  
; APPLICANT: AZIMZAL, Yalda  
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS  
; FILE REFERENCE: PF-0592-1 DIV  
; CURRENT APPLICATION NUMBER: US/10/427,631  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: US 09/786,240  
; PRIOR FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: PCT/US99/20989  
; PRIOR FILING DATE: 1999-09-09  
; PRIOR APPLICATION NUMBER: US 60/172,220  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: US 60/155,248  
; PRIOR FILING DATE: 1998-11-04  
; PRIOR APPLICATION NUMBER: US 60/133,642  
; PRIOR FILING DATE: 1999-05-11  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PERL Program  
; SEQ ID NO 7  
; LENGTH: 284  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: 1784742CD1  
US-10-427-631-7

Query Match 100.0%; Score 1530; DB 34; Length 284;  
Best Local Similarity 100.0%; Pred. No. 1.4e-154;  
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAESEAETPTGPEFSKYFPHGVRLPPFCRGKMEIANFPVRPSDVWIVTPKSGTSL 60  
Db 1 MAESEAETPTGPEFSKYFPHGVRLPPFCRGKMEIANFPVRPSDVWIVTPKSGTSL 60  
Qy 61 LQEVVYLVSGADPDEIGLWNIDQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Db 61 LQEVVYLVSGADPDEIGLWNIDQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Qy 121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFEHVQ 180  
Db 121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFEHVQ 180  
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAETHCHQLVDQCCNA 240  
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAETHCHQLVDQCCNA 240  
Qy 241 EALPVGRGRVGLWKDIFTVSNKEKFDLVYKQMGKCDLTTFDFYL 284  
Db 241 EALPVGRGRVGLWKDIFTVSNKEKFDLVYKQMGKCDLTTFDFYL 284

## RESULT 22

US-10-573-989-17  
; Sequence 17, Application US/10573989  
; GENERAL INFORMATION:  
; APPLICANT: Evotec NeuroSciences GmbH  
; APPLICANT: Von Der Kammer, Heinz  
; APPLICANT: Pohnner, Johannes  
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC USE OF A SULFOTRANSFERASE  
; TITLE OF INVENTION: FOR NEURODEGENERATIVE DISEASES  
; FILE REFERENCE: 2335.0140000/SRL/KPQ  
; CURRENT APPLICATION NUMBER: US/10/573,989  
; CURRENT FILING DATE: 2006-03-30  
; PRIOR APPLICATION NUMBER: PCT/EP2004/052353  
; PRIOR FILING DATE: 2004-09-29

; PRIOR APPLICATION NUMBER: 60/506,775  
; PRIOR FILING DATE: 2003-09-30  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 284  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-573-989-17

Query Match 100.0%; Score 1530; DB 35; Length 284;  
Best Local Similarity 100.0%; Pred. No. 1.4e-154;  
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAESEAETPTGPEFSKYFPHGVRLPPFCRGKMEIANFPVRPSDVWIVTPKSGTSL 60  
Db 1 MAESEAETPTGPEFSKYFPHGVRLPPFCRGKMEIANFPVRPSDVWIVTPKSGTSL 60  
Qy 61 LQEVVYLVSGADPDEIGLWNIDQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Db 61 LQEVVYLVSGADPDEIGLWNIDQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Qy 121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFEHVQ 180  
Db 121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFEHVQ 180  
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAETHCHQLVDQCCNA 240  
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAETHCHQLVDQCCNA 240  
Qy 241 EALPVGRGRVGLWKDIFTVSNKEKFDLVYKQMGKCDLTTFDFYL 284  
Db 241 EALPVGRGRVGLWKDIFTVSNKEKFDLVYKQMGKCDLTTFDFYL 284

## RESULT 23

US-10-757-262-110  
; Sequence 110, Application US/10757262  
; GENERAL INFORMATION:  
; APPLICANT: Karicheti, Venkateswarlu  
; APPLICANT: Silos-Santiago, Inmaculada  
; APPLICANT: Eliasof, Scott D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING  
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,  
; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,  
; TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43239, 44373, 51164,  
; TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,  
; TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,  
; TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,  
; TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,  
; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR  
; FILE REFERENCE: MPI03-007P1RNMNIM  
; CURRENT APPLICATION NUMBER: US/10/757,262  
; CURRENT FILING DATE: 2004-01-14  
; PRIOR APPLICATION NUMBER: US 60/440,318  
; PRIOR FILING DATE: 2003-01-15  
; PRIOR APPLICATION NUMBER: US 60/444,783  
; PRIOR FILING DATE: 2003-02-04  
; PRIOR APPLICATION NUMBER: US 60/457,901  
; PRIOR FILING DATE: 2003-03-27  
; PRIOR APPLICATION NUMBER: US 60/468,775  
; PRIOR FILING DATE: 2003-05-08  
; PRIOR APPLICATION NUMBER: US 60/471,614  
; PRIOR FILING DATE: 2003-05-19  
; PRIOR APPLICATION NUMBER: US 60/478,742  
; PRIOR FILING DATE: 2003-06-16  
; PRIOR APPLICATION NUMBER: US 60/488,529  
; PRIOR FILING DATE: 2003-07-18  
; PRIOR APPLICATION NUMBER: US 60/491,156  
; PRIOR FILING DATE: 2003-07-30  
; PRIOR APPLICATION NUMBER: US 60/499,594  
; PRIOR FILING DATE: 2003-09-02



; PRIOR APPLICATION NUMBER: US 60/506,332
; PRIOR FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 110
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-757-262-110

Query Match 100.0%; Score 1530; DB 37; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEAETPTGGEFESKYFEFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAETPTGGEFESKYFEFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60

QY 61 LOEVVYLVVSQADPDEIGLNMNIDEQLPVLEYPOQGLDIIKELTSRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVVSQADPDEIGLNMNIDEQLPVLEYPOQGLDIIKELTSRLIKSHLPYRFLPSD 120

QY 121 LHNGDSKVIYMARNP KD LVVSYQPHRSRLTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYMARNP KD LVVSYQPHRSRLTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

QY 181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240

QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284

RESULT 24
US-10-768-158-2
; Sequence 2, Application US/10768158
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: Karicheti, Venkateswarlu
; APPLICANT: Eliasof, Scott D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 16386, 15402, 21165, 1423,
; TITLE OF INVENTION: 636, 12303, 21425, 27410, 38554, 38555, 55063, 57145, 59914,
; TITLE OF INVENTION: 94921, 16852, 33260, 58573, 30911, 85913, 14303, 16816,
; TITLE OF INVENTION: 17827 OR 32620
; FILE REFERENCE: MPI03-012P1RNMNIM
; CURRENT APPLICATION NUMBER: US/10/768,158
; PRIOR FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: US 60/444,781
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/452,291
; PRIOR FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: US 60/454,540
; PRIOR FILING DATE: 2003-03-13
; PRIOR APPLICATION NUMBER: US 60/478,805
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/491,048
; PRIOR FILING DATE: 2003-07-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-768-158-2

Query Match 100.0%; Score 1530; DB 37; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEAETPTGGEFESKYFEFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAETPTGGEFESKYFEFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60

QY 61 LOEVVYLVVSQADPDEIGLNMNIDEQLPVLEYPOQGLDIIKELTSRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVVSQADPDEIGLNMNIDEQLPVLEYPOQGLDIIKELTSRLIKSHLPYRFLPSD 120

QY 121 LHNGDSKVIYMARNP KD LVVSYQPHRSRLTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYMARNP KD LVVSYQPHRSRLTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

QY 181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240

QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284

RESULT 25
US-10-932-348-408
; Sequence 408, Application US/10932348
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN DRUG-METABOLIZING ENZYMES, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001481
; CURRENT APPLICATION NUMBER: US/10/932,348
; NUMBER OF SEQ ID NOS: 30431
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 408
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-932-348-408

Query Match 100.0%; Score 1530; DB 39; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEAETPTGGEFESKYFEFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAETPTGGEFESKYFEFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60

QY 61 LOEVVYLVVSQADPDEIGLNMNIDEQLPVLEYPOQGLDIIKELTSRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVVSQADPDEIGLNMNIDEQLPVLEYPOQGLDIIKELTSRLIKSHLPYRFLPSD 120

QY 121 LHNGDSKVIYMARNP KD LVVSYQPHRSRLTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYMARNP KD LVVSYQPHRSRLTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

QY 181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240

QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284

RESULT 26
US-10-990-328-7686
; Sequence 7686, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF

```
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7686
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328-7686

Query Match      100.0%; Score 1530; DB 39; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEATPTSGEFESKYFEFGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Db      1 MAESEATPTSGEFESKYFEFGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60

Qy      61 LQEVVYLVSQADPDEIGLNMIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LQEVVYLVSQADPDEIGLNMIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYARNPKDLVSYQFHRSLRTMSYRGTFQFCRRFMDKLGYSWFEHVQ 180
Db      121 LHNGDSKVIYARNPKDLVSYQFHRSLRTMSYRGTFQFCRRFMDKLGYSWFEHVQ 180

Qy      181 EFWEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCNA 240
Db      181 EFWEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCNA 240

Qy      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284
Db      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284

RESULT 27
US-10-990-328A-7686
; Sequence 7686, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7686
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328A-7686

Query Match      100.0%; Score 1530; DB 39; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEATPTSGEFESKYFEFGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Db      1 MAESEATPTSGEFESKYFEFGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60

Qy      61 LQEVVYLVSQADPDEIGLNMIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LQEVVYLVSQADPDEIGLNMIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYARNPKDLVSYQFHRSLRTMSYRGTFQFCRRFMDKLGYSWFEHVQ 180
Db      121 LHNGDSKVIYARNPKDLVSYQFHRSLRTMSYRGTFQFCRRFMDKLGYSWFEHVQ 180

Qy      181 EFWEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCNA 240
Db      181 EFWEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCNA 240

Qy      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284
Db      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284

RESULT 28
US-11-154-298-60
; Sequence 60, Application US/11154298
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.01
; CURRENT APPLICATION NUMBER: US/11/154,298
; CURRENT FILING DATE: 2005-06-15
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-154-298-60

Query Match      100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAESEATPTSGEFESKYFEFGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Db      1 MAESEATPTSGEFESKYFEFGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60

Qy      61 LQEVVYLVSQADPDEIGLNMIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db      61 LQEVVYLVSQADPDEIGLNMIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy      121 LHNGDSKVIYARNPKDLVSYQFHRSLRTMSYRGTFQFCRRFMDKLGYSWFEHVQ 180
Db      121 LHNGDSKVIYARNPKDLVSYQFHRSLRTMSYRGTFQFCRRFMDKLGYSWFEHVQ 180

Qy      181 EFWEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCNA 240
Db      181 EFWEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCNA 240

Qy      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284
Db      241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTTFDFYL 284

RESULT 29
US-11-154-298-81
; Sequence 81, Application US/11154298
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
```

```
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.01
; CURRENT APPLICATION NUMBER: US/11/154,298
; CURRENT FILING DATE: 2005-06-15
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-154-298-81

Query Match      100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEAETPTGGEFESKYFEFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60
DB 1 MAESEAETPTGGEFESKYFEFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60
QY 61 LQEVVYLSQAGADPEIGLNMNIDEQLPVLEYPOGLDIIKELTSRLIKSHLPYRFLPSD 120
DB 61 LQEVVYLSQAGADPEIGLNMNIDEQLPVLEYPOGLDIIKELTSRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNPDLVSYQQFHRSLRTMSYRGTFQEFCCRFRMNDKLGYSWFHVQ 180
DB 121 LHNGDSKVIYMARNPDLVSYQQFHRSLRTMSYRGTFQEFCCRFRMNDKLGYSWFHVQ 180
QY 181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLTEALTECHQLVDQCCNA 240
DB 181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLTEALTECHQLVDQCCNA 240
QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
DB 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 31
US-11-154-298-104
; Sequence 104, Application US/11154298
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.01
; CURRENT APPLICATION NUMBER: US/11/154,298
; CURRENT FILING DATE: 2005-06-15
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-154-298-104

Query Match      100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEAETPTGGEFESKYFEFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60
DB 1 MAESEAETPTGGEFESKYFEFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60
QY 61 LQEVVYLSQAGADPEIGLNMNIDEQLPVLEYPOGLDIIKELTSRLIKSHLPYRFLPSD 120
DB 61 LQEVVYLSQAGADPEIGLNMNIDEQLPVLEYPOGLDIIKELTSRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNPDLVSYQQFHRSLRTMSYRGTFQEFCCRFRMNDKLGYSWFHVQ 180
DB 121 LHNGDSKVIYMARNPDLVSYQQFHRSLRTMSYRGTFQEFCCRFRMNDKLGYSWFHVQ 180
QY 181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLTEALTECHQLVDQCCNA 240
DB 181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLTEALTECHQLVDQCCNA 240
QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
DB 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 30
US-11-154-298-103
; Sequence 103, Application US/11154298
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.01
; CURRENT APPLICATION NUMBER: US/11/154,298
; CURRENT FILING DATE: 2005-06-15
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
```

```
Db 61 LQEVVYLVSGQADPDEIGLNMIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Qy 181 EFWHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db 181 EFWHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 32
US-11-154-298-708
; Sequence 708, Application US/11154298
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.01
; CURRENT APPLICATION NUMBER: US/11/154,298
; CURRENT FILING DATE: 2005-06-15
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 708
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-154-298-708

Query Match 100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEATPSTPGFESKYFEGHVRLLPPFCRGKMEIEANFPVRPSDVIIVTPKSGTSL 60
Db 1 MAESEATPSTPGFESKYFEGHVRLLPPFCRGKMEIEANFPVRPSDVIIVTPKSGTSL 60
Qy 61 LQEVVYLVSGQADPDEIGLNMIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLVSGQADPDEIGLNMIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Qy 181 EFWHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db 181 EFWHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 34
US-11-155-906-81
; Sequence 81, Application US/11155906
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.02
; CURRENT APPLICATION NUMBER: US/11/155,906
; CURRENT FILING DATE: 2005-06-17
```

```
RESULT 33
US-11-155-906-60
; Sequence 60, Application US/11155906
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.02
; CURRENT APPLICATION NUMBER: US/11/155,906
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-155-906-60

Query Match 100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEATPSTPGFESKYFEGHVRLLPPFCRGKMEIEANFPVRPSDVIIVTPKSGTSL 60
Db 1 MAESEATPSTPGFESKYFEGHVRLLPPFCRGKMEIEANFPVRPSDVIIVTPKSGTSL 60
Qy 61 LQEVVYLVSGQADPDEIGLNMIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLVSGQADPDEIGLNMIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Qy 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Qy 181 EFWHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Db 181 EFWHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALTEHCHQLVDQCCNA 240
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 34
US-11-155-906-81
; Sequence 81, Application US/11155906
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.02
; CURRENT APPLICATION NUMBER: US/11/155,906
; CURRENT FILING DATE: 2005-06-17
```

Query Match	100.0%;	Score 1530;	DB 41;	Length 284;																													
Best Local Similarity	100.0%;	Pred. No. 1.4e-154;																															
Matches 284;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;																													
Qy	1	MAESEA	TP	TPGPFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60																													
Db	1	MAESEA	TP	TPGPFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60																													
Qy	61	LQEVVYL	VSQ	ADPEIGLNNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120																													
Db	61	LQEVVYL	VSQ	ADPEIGLNNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120																													
Qy	121	LHNGDS	KVI	YMARNP	KDLV	SSYQF	HRSLR	TMSY	RGTF	QECRR	PMND	KL	GYG	SW	FHVQ 180																		
Db	121	LHNGDS	KVI	YMARNP	KDLV	SSYQF	HRSLR	TMSY	RGTF	QECRR	PMND	KL	GYG	SW	FHVQ 180																		
Qy	181	EFWEHR	MDS	NNVLF	KYED	MHRDLVT	WV	EQ	LARF	LG	VS	CD	K	AQ	LEALTEHCHOLVDOCCNA 240																		
Db	181	EFWEHR	MDS	NNVLF	KYED	MHRDLVT	WV	EQ	LARF	LG	VS	CD	K	AQ	LEALTEHCHOLVDOCCNA 240																		
Qy	241	EALPVG	RGR	VG	LWK	DI	F	T	V	S	M	N	E	K	F	D	L	V	Y	K	Q	M	G	K	C	D	L	T	F	D	F	Y	L 284
Db	241	EALPVG	RGR	VG	LWK	DI	F	T	V	S	M	N	E	K	F	D	L	V	Y	K	Q	M	G	K	C	D	L	T	F	D	F	Y	L 284
RESULT 36																																	
US-11-155-906-104																																	
; Sequence 104, Application US/11155906																																	
; GENERAL INFORMATION:																																	
; APPLICANT: Jakobovits, Ava																																	
; APPLICANT: Challita-Eid, Pia M.																																	
; APPLICANT: Paris, Mary																																	
; APPLICANT: Ge, Wangmao																																	
; APPLICANT: Hubert, Rene S.																																	
; APPLICANT: Morrison, Karen Jane Meyrick																																	
; APPLICANT: Morrison, Robert Kendall																																	
; APPLICANT: Raitano, Arthur B.																																	
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE																																	
; FILE REFERENCE: 51158-20040.02																																	
; CURRENT APPLICATION NUMBER: US/11/155.906																																	
; CURRENT FILING DATE: 2005-06-17																																	
; PRIOR APPLICATION NUMBER: US10/121,024																																	
; PRIOR FILING DATE: 2002-04-10																																	
; PRIOR APPLICATION NUMBER: US60/283,112																																	
; PRIOR FILING DATE: 2001-04-10																																	
; PRIOR APPLICATION NUMBER: US60/282,739																																	
; PRIOR FILING DATE: 2001-04-10																																	
; PRIOR APPLICATION NUMBER: US60/286,630																																	
; PRIOR FILING DATE: 2001-04-25																																	
; NUMBER OF SEQ ID NOS: 708																																	
; SOFTWARE: Fast-Seq for Windows Version 4.0																																	
; SEQ ID NO 104																																	
; LENGTH: 284																																	
; TYPE: PRT																																	
; ORGANISM: Homo sapiens																																	
US-11-155-906-104																																	
Query Match	100.0%;	Score 1530;	DB 41;	Length 284;																													
Best Local Similarity	100.0%;	Pred. No. 1.4e-154;																															
Matches 284;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;																													
Qy	1	MAESEA	TP	TPGPFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60																													
Db	1	MAESEA	TP	TPGPFESKYFEFHGVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60																													
Qy	61	LQEVVYL	VSQ	ADPEIGLNNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120																													
Db	61	LQEVVYL	VSQ	ADPEIGLNNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120																													
Qy	121	LHNGDS	KVI	YMARNP	KDLV	SSYQF	HRSLR	TMSY	RGTF	QECRR	PMND	KL	GYG	SW	FHVQ 180																		
Db	121	LHNGDS	KVI	YMARNP	KDLV	SSYQF	HRSLR	TMSY	RGTF	QECRR	PMND	KL	GYG	SW	FHVQ 180																		

```

QY 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

QY 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTDFDYL 284
Db 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTDFDYL 284

RESULT 37
US-11-155-906-708
; Sequence 708, Application US/11155906
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.02
; CURRENT APPLICATION NUMBER: US/11/155,906
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 708
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-155-906-708

Query Match 100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEATPSTPGFEFSKYFEFHGVRLPPFCRGKMEIANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPSTPGFEFSKYFEFHGVRLPPFCRGKMEIANFPVRPSDVWIVTPKSGTSL 60

QY 61 LOEVVYLVSQGADPDEIGLMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQGADPDEIGLMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

QY 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMDKLGYSWFEHVQ 180
Db 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMDKLGYSWFEHVQ 180

QY 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

QY 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTDFDYL 284
Db 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTDFDYL 284

RESULT 38
US-11-156-231-60
; Sequence 60, Application US/11156231
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.03
; CURRENT APPLICATION NUMBER: US/11/156,231
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630

```

```

; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.03
; CURRENT APPLICATION NUMBER: US/11/156,231
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-156-231-60

Query Match 100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEATPSTPGFEFSKYFEFHGVRLPPFCRGKMEIANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPSTPGFEFSKYFEFHGVRLPPFCRGKMEIANFPVRPSDVWIVTPKSGTSL 60

QY 61 LOEVVYLVSQGADPDEIGLMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQGADPDEIGLMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

QY 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMDKLGYSWFEHVQ 180
Db 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMDKLGYSWFEHVQ 180

QY 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLALETCHQLVDQCCNA 240

QY 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTDFDYL 284
Db 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTDFDYL 284

RESULT 39
US-11-156-231-81
; Sequence 81, Application US/11156231
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.03
; CURRENT APPLICATION NUMBER: US/11/156,231
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630

```

```
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-156-231-81

Query Match      100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEAETPSTPGFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
DB 1 MAESEAETPSTPGFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60

QY 61 LQEVVYLVSQGADPDEIGLMNIDEQLPVLEYQPGLDIIKELTSRLIKSHLPYRFLPSD 120
DB 61 LQEVVYLVSQGADPDEIGLMNIDEQLPVLEYQPGLDIIKELTSRLIKSHLPYRFLPSD 120

QY 121 LHNGDSKVIYMARNP KD LVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
DB 121 LHNGDSKVIYMARNP KD LVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

QY 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAETHCHQLVDQCCNA 240
DB 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAETHCHQLVDQCCNA 240

QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
DB 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 40
US-11-156-231-103
; Sequence 103, Application US/11156231
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Ava
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.03
; CURRENT APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-156-231-104

Query Match      100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEAETPSTPGFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
DB 1 MAESEAETPSTPGFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60

QY 61 LQEVVYLVSQGADPDEIGLMNIDEQLPVLEYQPGLDIIKELTSRLIKSHLPYRFLPSD 120
DB 61 LQEVVYLVSQGADPDEIGLMNIDEQLPVLEYQPGLDIIKELTSRLIKSHLPYRFLPSD 120

QY 121 LHNGDSKVIYMARNP KD LVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
DB 121 LHNGDSKVIYMARNP KD LVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

QY 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAETHCHQLVDQCCNA 240
DB 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAETHCHQLVDQCCNA 240

QY 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
DB 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 40
US-11-156-231-103
; Sequence 103, Application US/11156231
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Ava
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.03
; CURRENT APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-156-231-103

Query Match      100.0%; Score 1530; DB 41; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.4e-154;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAESEAETPSTPGFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
DB 1 MAESEAETPSTPGFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
```

RESULT 42

```

US-11-156-231-708
; Sequence 708, Application US/11156231
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING A
; FILE REFERENCE: 51158-20040:03
; CURRENT APPLICATION NUMBER: US/11156,231
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 708
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-156-231-708

```

Query Match	100.0%	Score 1530;	DB 41;	Length 284;
Best Local Similarity	100.0%	Pred. No. 1.4e-154;		
Matches 284;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAESEAETPSTPGSEFFESKYFEFHGVRUPPFRCGRMEEIANFPVRPSDVMIITYPKSGTSL	60	
Db	1	MAESEAETPSTPGSEFFESKYFEFHGVRUPPFRCGRMEEIANFPVRPSDVMIITYPKSGTSL	60	
Qy	61	LOEVVYLVQSOGADPDETLGMINIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPFRFLPSD	120	
Db	61	LOEVVYLVQSOGADPDETLGMINIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPFRFLPSD	120	
Qy	121	LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMDNKLGYGSWFHFVQ	180	
Db	121	LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMDNKLGYGSWFHFVQ	180	
Qy	181	EFWEHRMDSNVLFUKYEDMHRDLVTMVQEQIARFLGVSCDKAQLEALTEHCHQLVDQCCNA	240	
Db	181	EFWEHRMDSNVLFUKYEDMHRDLVTMVQEQIARFLGVSCDKAQLEALTEHCHQLVDQCCNA	240	
Qy	241	EALPVGRRGVLGKWDIEFTVSNNEKFDLVYKQMGKCDLTFDFYL	284	
Db	241	EALPVGRRGVLGKWDIEFTVSNNEKFDLVYKQMGKCDLTFDFYL	284	

## RESULT 43

```

US-60-444-781-2
; Sequence 2, Application US/60444781
; GENERAL INFORMATION:
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: Methods and compositions in treating
; TITLE OF INVENTION: pain and painful disorders using 16386, 15402, 21165 or 1423
; FILE REFERENCE: MP103-012P1M
; CURRENT APPLICATION NUMBER: US/60/444,781
; CURRENT FILING DATE: 2003-02-04
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 284

```

```

; TYPE: PRT
; ORGANISM: homosapien
US-60-444-781-2

```

Query Match	100.0%;	Score 1530;	DB 49;	Length 284;
Best Local Similarity	100.0%;	Pred. NO. 1.4e-154;		
Matches 284;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAESEAETPTGDEFESKYFEFHGVRUPPFCRGMEEIANFPVRPSDVMIYTPKSGTSL	60	
Db	1	MAESEAETPTGDEFESKYFEFHGVRUPPFCRGMEEIANFPVRPSDVMIYTPKSGTSL	60	
Qy	61	LOEVVYLVSOGADPELGLMNIIDQLVLEVPQGLDIIKELTSPRLIKSHLPFRFPLPSD	120	
Db	61	LOEVVYLVSOGADPELGLMNIIDQLVLEVPQGLDIIKELTSPRLIKSHLPFRFPLPSD	120	
Qy	121	LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQBFCCRPMNDKLGYSWFPHVQ	180	
Db	121	LHNGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQBFCCRPMNDKLGYSWFPHVQ	180	
Qy	181	EFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQALTEHCHQLVDQCCNA	240	
Db	181	EFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQALTEHCHQLVDQCCNA	240	
Qy	241	EALPVGGRGVLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL	284	
Db	241	EALPVGGRGVLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL	284	

RESULT 44

```

US-60-491-156-24
; Sequence 24, Application US/60491156
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Karichetti, Venkateswarlu
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Eliasof, Scott
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 14303, 16816, 17
; TITLE OF INVENTION: 619, 1423, 2159, 8263, 15402, 16209, 16386,
; TITLE OF INVENTION: 41897
; FILE REFERENCE: MPI03-098P1M
; CURRENT APPLICATION NUMBER: US/60/491,156
; CURRENT FILING DATE: 2003-07-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-60-491-156-24

```

Query Match	100.0%	Score 1530;	DB 49;	Length 284;
Best Local Similarity	100.0%;	Pred. No. 1.4e-154;		
Matches 284;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAESEAETPSTPGFESKYFEFHGVRLLPPFCRGKMEBIANPPVRPSDVMIYTPKSGTSL	60	
DB	1	MAESEAETPSTPGFESKYFEFHGVRLLPPFCRGKMEBIANPPVRPSDVMIYTPKSGTSL	60	
QY	61	LOEVVYLVSQGADPDEITGLMNIIDQLPVLEYQPGLDIIKELTSPLRIKSHLPYRFLPSD	120	
DB	61	LOEVVYLVSQGADPDEITGLMNIIDQLPVLEYQPGLDIIKELTSPLRIKSHLPYRFLPSD	120	
QY	121	LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQFECRRFMDNKLGYSGSWFHVQ	180	
DB	121	LHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSYRGTFQFECRRFMDNKLGYSGSWFHVQ	180	
QY	181	EFWEHRMDSNVLFUKYEDMHRDLVTWVEQLARFLGVSCDKAKQALTEHCHQLVDQCNA	240	
DB	181	EFWEHRMDSNVLFUKYEDMHRDLVTWVEQLARFLGVSCDKAKQALTEHCHQLVDQCNA	240	
QY	241	EALPVGGRGRVGLWKXDIETVSNMEXFDLVYKQKMGKCDLTFDFYL	284	



Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284  
|||||  
RESULT 45  
US-60-499-964-408  
; Sequence 408, Application US/60499964  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES  
; TITLE OF INVENTION: ENCODING HUMAN DRUG-METABOLIZING ENZYMES, METHODS OF  
; TITLE OF INVENTION: DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001481  
; CURRENT APPLICATION NUMBER: US/60/499,964  
; CURRENT FILING DATE: 2003-09-04  
; NUMBER OF SEQ ID NOS: 30431  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 408  
; LENGTH: 284  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-499-964-408  
Query Match 100.0%; Score 1530; DB 49; Length 284;  
Best Local Similarity 100.0%; Pred. No. 1.4e-154; Mismatches 0; Indels 0; Gaps 0;  
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAESEAEPTSPGPEESKYFFHGVRLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60  
Db 1 MAESEAEPTSPGPEESKYFFHGVRLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60  
Qy 61 LOEVVYLVVSQAGADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Db 61 LOEVVYLVVSQAGADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Qy 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFEHVQ 180  
Db 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFEHVQ 180  
Qy 191 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 240  
Db 191 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 240  
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284  
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284  
|||||  
RESULT 46  
US-10-170-205E-16293  
; Sequence 16293, Application US/10170205E  
; GENERAL INFORMATION:  
; APPLICANT: ADAMS, Mark  
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN  
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF  
; FILE REFERENCE: CL001381  
; CURRENT APPLICATION NUMBER: US/10/170,205E  
; CURRENT FILING DATE: 2002-06-13  
; NUMBER OF SEQ ID NOS: 40312  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 16293  
; LENGTH: 311  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-170-205E-16293  
Query Match 100.0%; Score 1530; DB 31; Length 311;  
Best Local Similarity 100.0%; Pred. No. 1.6e-154; Mismatches 0; Indels 0; Gaps 0;  
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAESEAEPTSPGPEESKYFFHGVRLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60  
Db 28 MAESEAEPTSPGPEESKYFFHGVRLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 87

Qy 61 LOEVVYLVVSQAGADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Db 88 LOEVVYLVVSQAGADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 147  
Qy 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFEHVQ 180  
Db 148 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFEHVQ 207  
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 240  
Db 208 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 267  
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284  
Db 268 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 311  
|||||  
RESULT 47  
US-60-245-241-166  
; Sequence 166, Application US/60245241  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET ENZYMES,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET ENZYMES,  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: CL000888  
; CURRENT APPLICATION NUMBER: US/60/245,241  
; CURRENT FILING DATE: 2000-11-03  
; NUMBER OF SEQ ID NOS: 255  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 166  
; LENGTH: 311  
; TYPE: PRT  
; ORGANISM: Human  
US-60-245-241-166  
Query Match 100.0%; Score 1530; DB 47; Length 311;  
Best Local Similarity 100.0%; Pred. No. 1.6e-154; Mismatches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAESEAEPTSPGPEESKYFFHGVRLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60  
Db 28 MAESEAEPTSPGPEESKYFFHGVRLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 87  
Qy 61 LOEVVYLVVSQAGADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Db 88 LOEVVYLVVSQAGADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 147  
Qy 121 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFEHVQ 180  
Db 148 LHNGDSKVIYMARNP KDVLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFEHVQ 207  
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 240  
Db 208 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLVDQCCNA 267  
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284  
Db 268 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 311  
|||||  
RESULT 48  
US-60-258-016-46  
; Sequence 46, Application US/60258016  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET ENZYMES,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET ENZYMES,  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: CL001037-PROV  
; CURRENT APPLICATION NUMBER: US/60/258,016  
; CURRENT FILING DATE: 2000-12-27

; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 46  
; LENGTH: 311  
; TYPE: PRT  
; ORGANISM: Human  
US-60-258-016-46

Query Match 100.0%; Score 1530; DB 47; Length 311;  
Best Local Similarity 100.0%; Pred. No. 1.6e-154;  
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESAETPTGPEFESKYFEHGVRLPPFCRGKMEETANFPVRPSDVMIYTPKSGTSL 60  
Db 28 MAESAETPTGPEFESKYFEHGVRLPPFCRGKMEETANFPVRPSDVMIYTPKSGTSL 87  
Qy 61 LQEVVYLSQAGDPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Db 88 LQEVVYLSQAGDPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 147  
Qy 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180  
Db 148 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 207  
Qy 181 EFWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCKAQLALETCHQLVDQCCNA 240  
Db 208 EFWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCKAQLALETCHQLVDQCCNA 267  
Qy 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFPYL 284  
Db 268 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFPYL 311

## RESULT 49

US-09-791-537-89260  
; Sequence 89260, Application US/09791537  
; GENERAL INFORMATION:

; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 89260  
; LENGTH: 284  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-89260

Query Match 98.4%; Score 1506; DB 27; Length 284;  
Best Local Similarity 97.9%; Pred. No. 5.3e-152;  
Matches 278; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAESAETPTGPEFESKYFEHGVRLPPFCRGKMEETANFPVRPSDVMIYTPKSGTSL 60  
Db 1 MAESAETPTGPEFESKYFEHGVRLPPFCRGKMEETANFPVRPSDVMIYTPKSGTSL 60  
Qy 61 LQEVVYLSQAGDPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Db 61 LQEVVYLSQAGDPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Qy 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180  
Db 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180  
Qy 181 EFWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCKAQLALETCHQLVDQCCNA 240  
Db 181 EFWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCKAQLALETCHQLVDQCCNA 240

Qy 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFPYL 284  
Db 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFPYL 284

## RESULT 50

US-10-205-331-46  
; Sequence 46, Application US/10205331  
; GENERAL INFORMATION:

; APPLICANT: Warner-Lambert Company  
; APPLICANT: Lee, Kevin  
; APPLICANT: Dixon, Alistair  
; APPLICANT: Brooksbank, Robert  
; APPLICANT: Pinnock, Robert  
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain  
; FILE REFERENCE: WL-A-018199  
; CURRENT APPLICATION NUMBER: US/10/205,331  
; CURRENT FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: GB 0118354.0  
; PRIOR FILING DATE: 2001-07-27  
; NUMBER OF SEQ ID NOS: 117  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 46  
; LENGTH: 284  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; OTHER INFORMATION: Sulfotransferase-like protein  
US-10-205-331-46

Query Match 98.4%; Score 1506; DB 32; Length 284;  
Best Local Similarity 97.9%; Pred. No. 5.3e-152;  
Matches 278; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAESAETPTGPEFESKYFEHGVRLPPFCRGKMEETANFPVRPSDVMIYTPKSGTSL 60  
Db 1 MAESAETPTGPEFESKYFEHGVRLPPFCRGKMEETANFPVRPSDVMIYTPKSGTSL 60  
Qy 61 LQEVVYLSQAGDPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Db 61 LQEVVYLSQAGDPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Qy 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180  
Db 121 LHNGDSKVIYARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180  
Qy 181 EFWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCKAQLALETCHQLVDQCCNA 240  
Db 181 EFWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCKAQLALETCHQLVDQCCNA 240  
Qy 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFPYL 284  
Db 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFPYL 284

## RESULT 51

US-60-229-512-225

; Sequence 225, Application US/60229512  
; GENERAL INFORMATION:

; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN PHASE II  
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, NUCLEIC ACID MOLECULES ENCODING  
; TITLE OF INVENTION: HUMAN DRUG-METABOLIZING PHASE II PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CLO00770  
; CURRENT APPLICATION NUMBER: US/60/229,512  
; CURRENT FILING DATE: 2000-09-05  
; NUMBER OF SEQ ID NOS: 516  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 225  
; LENGTH: 319  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-229-512-225

Query Match 98.2%; Score 1502.5; DB 47; Length 319;  
Best Local Similarity 89.0%; Pred. No. 1.5e-151;  
Matches 284; Conservative 0; Mismatches 0; Indels 35; Gaps 1;  
  
QY 1 MAESEAETPSTGPEFESKYFEHGVRLPPFCRGMEEIANFPVRPSDVWIVTPKSGTSL 56  
DB 1 MAESEAETPSTGPEFESKYFEHGVRLPPFCRGMEEIANFPVRPSDVWIVTPKSGRHH 60  
  
QY 57 -----GTSLLQEVVYLVSOQADPDEIGLWNIDQ 85  
DB 61 AGRDEEPAETGAADMGKQWVKLCPFLLGTSLLQEVVYLVSOQADPDEIGLWNIDQ 120  
  
QY 86 LPVLYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDLVVSYYOF 145  
DB 121 LPVLYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDLVVSYYOF 180  
  
QY 146 HRSRTMSYRGTFQFCRRFMNDKLGYSWFEHGVRLPPFCRGMEEIANFPVRPSDVWIVTPKSGTSL 205  
DB 181 HRSRTMSYRGTFQFCRRFMNDKLGYSWFEHGVRLPPFCRGMEEIANFPVRPSDVWIVTPKSGTSL 240  
  
QY 206 MVEQLARFLGVSCKAQAQLEALTEHCHQVLDQCNAEALPVGRVGLWKDIFTVSMNEKF 265  
DB 241 MVEQLARFLGVSCKAQAQLEALTEHCHQVLDQCNAEALPVGRVGLWKDIFTVSMNEKF 300  
  
QY 266 DLVYKQKMGKCDLTDFYLL 284  
DB 301 DLVYKQKMGKCDLTDFYLL 319

RESULT 52  
US-10-170-205E-27511  
; Sequence 27511, Application US/10170205E  
; GENERAL INFORMATION:  
; APPLICANT: ADAMS, Mark  
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN  
; FILE REFERENCE: CAPTURE AGENTS, AND USES THEREOF  
; CURRENT APPLICATION NUMBER: US/10/170,205E  
; CURRENT FILING DATE: 2002-06-13  
; NUMBER OF SEQ ID NOS: 40312  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 27511  
; LENGTH: 260  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-170-205E-27511

Query Match 87.1%; Score 1332; DB 31; Length 260;  
Best Local Similarity 96.9%; Pred. No. 2.2e-133;  
Matches 249; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
  
QY 1 MAESEAETPSTGPEFESKYFEHGVRLPPFCRGMEEIANFPVRPSDVWIVTPKSGTSL 60  
DB 1 MAESEAETPSTGPEFESKYFEHGVRLPPFCRGMEEIANFPVRPSDVWIVTPKSGTSL 60  
  
QY 61 LOEVVYLVSOQADPDEIGLWNIDQPLVLEYVQPGDLIIKELTSPRLIKSHLPYRFLPSD 120  
DB 61 LOEVVYLVSOQADPDEIGLWNIDQPLVLEYVQPGDLIIKELTSPRLIKSHLPYRFLPSD 120  
  
QY 121 LHNGDSKVIYMARNPDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHGV 180  
DB 121 LHNGDSKVIYMARNPDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHGV 180  
  
QY 181 EFWEHRMDSNVLFUKYEDMHRDLVTMVEQLARFLGVSCKAQAQLEALTEHCHQVLDQCNA 240  
DB 181 EFWEHRMDSNVLFUKYEDMHRDLVTMVEQLARFLGVSCKAQAQLEALTEHCHQVLDQCNA 240  
  
QY 241 EALPVGRVGLWKDIF 257  
DB 241 EALPVGRVGLWKDIF 257

RESULT 53  
US-10-932-348-406  
; Sequence 406, Application US/10932348  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES  
; TITLE OF INVENTION: ENCODING HUMAN DRUG-METABOLIZING ENZYMES, METHODS OF  
; FILE REFERENCE: CL001481  
; CURRENT APPLICATION NUMBER: US/10/932,348  
; CURRENT FILING DATE: 2004-09-02  
; NUMBER OF SEQ ID NOS: 30431  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 406  
; LENGTH: 260  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-932-348-406

Query Match 87.1%; Score 1332; DB 39; Length 260;  
Best Local Similarity 96.9%; Pred. No. 2.2e-133;  
Matches 249; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
  
QY 1 MAESEAETPSTGPEFESKYFEHGVRLPPFCRGMEEIANFPVRPSDVWIVTPKSGTSL 60  
DB 1 MAESEAETPSTGPEFESKYFEHGVRLPPFCRGMEEIANFPVRPSDVWIVTPKSGTSL 60  
  
QY 61 LOEVVYLVSOQADPDEIGLWNIDQPLVLEYVQPGDLIIKELTSPRLIKSHLPYRFLPSD 120  
DB 61 LOEVVYLVSOQADPDEIGLWNIDQPLVLEYVQPGDLIIKELTSPRLIKSHLPYRFLPSD 120  
  
QY 121 LHNGDSKVIYMARNPDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHGV 180  
DB 121 LHNGDSKVIYMARNPDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHGV 180  
  
QY 181 EFWEHRMDSNVLFUKYEDMHRDLVTMVEQLARFLGVSCKAQAQLEALTEHCHQVLDQCNA 240  
DB 181 EFWEHRMDSNVLFUKYEDMHRDLVTMVEQLARFLGVSCKAQAQLEALTEHCHQVLDQCNA 240  
  
QY 241 EALPVGRVGLWKDIF 257  
DB 241 EALPVGRVGLWKDIF 257

RESULT 54  
US-10-990-328-7684  
; Sequence 7684, Application US/10990328  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES  
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND  
; FILE REFERENCE: USES THEREOF  
; CURRENT APPLICATION NUMBER: US/10/990,328  
; CURRENT FILING DATE: 2004-11-17  
; NUMBER OF SEQ ID NOS: 55824  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7684  
; LENGTH: 260  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-990-328-7684

Query Match 87.1%; Score 1332; DB 39; Length 260;  
Best Local Similarity 96.9%; Pred. No. 2.2e-133;  
Matches 249; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
  
QY 1 MAESEAETPSTGPEFESKYFEHGVRLPPFCRGMEEIANFPVRPSDVWIVTPKSGTSL 60  
DB 1 MAESEAETPSTGPEFESKYFEHGVRLPPFCRGMEEIANFPVRPSDVWIVTPKSGTSL 60  
  
QY 61 LOEVVYLVSOQADPDEIGLWNIDQPLVLEYVQPGDLIIKELTSPRLIKSHLPYRFLPSD 120  
DB 61 LOEVVYLVSOQADPDEIGLWNIDQPLVLEYVQPGDLIIKELTSPRLIKSHLPYRFLPSD 120

```
Db 61 LOEVVYLVSQADPDEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
;
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-499-964-406

Query Match 87.1%; Score 1332; DB 39; Length 260;
Best Local Similarity 96.9%; Pred. No. 2.2e-133;
Matches 249; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MAESEATPTPGFESKYFEFHGVRLLPPFCRGMKEEIANPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPTPGFESKYFEFHGVRLLPPFCRGMKEEIANPVRPSDVWIVTPKSGTSL 60

Qy 61 LOEVVYLVSQADPDEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTMVQELARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTMVQELARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIF 257
Db 241 EALPVGRAHCVFARKIF 257

RESULT 55
US-10-990-328A-7684
; Sequence 7684, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558924
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7684
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328A-7684

Query Match 87.1%; Score 1332; DB 39; Length 260;
Best Local Similarity 96.9%; Pred. No. 2.2e-133;
Matches 249; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MAESEATPTPGFESKYFEFHGVRLLPPFCRGMKEEIANPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPTPGFESKYFEFHGVRLLPPFCRGMKEEIANPVRPSDVWIVTPKSGTSL 60

Qy 61 LOEVVYLVSQADPDEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTMVQELARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTMVQELARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIF 257
Db 241 EALPVGRAHCVFARKIF 257

RESULT 56
US-60-499-964-406
; Sequence 406, Application US/60499964
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN DRUG-METABOLIZING ENZYMES, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001481
; CURRENT APPLICATION NUMBER: US/60/499,964
; CURRENT FILING DATE: 2003-09-04
; NUMBER OF SEQ ID NOS: 30431
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 406
```

```
;
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-499-964-406

Query Match 87.1%; Score 1332; DB 49; Length 260;
Best Local Similarity 96.9%; Pred. No. 2.2e-133;
Matches 249; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MAESEATPTPGFESKYFEFHGVRLLPPFCRGMKEEIANPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPTPGFESKYFEFHGVRLLPPFCRGMKEEIANPVRPSDVWIVTPKSGTSL 60

Qy 61 LOEVVYLVSQADPDEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTMVQELARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTMVQELARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIF 257
Db 241 EALPVGRAHCVFARKIF 257

RESULT 57
US-10-932-348-407
; Sequence 407, Application US/10932348
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN DRUG-METABOLIZING ENZYMES, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001481
; CURRENT APPLICATION NUMBER: US/10/932,348
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 30431
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 407
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-932-348-407

Query Match 82.8%; Score 1267; DB 39; Length 286;
Best Local Similarity 84.5%; Pred. No. 2.4e-126;
Matches 246; Conservative 7; Mismatches 26; Indels 12; Gaps 3;

Qy 1 MAESEATPTPGFESKYFEFHGVRLLPPFCRGMKEEIANPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPTPGFESKYFEFHGVRLLPPFCRGMKEEIANPVRPSDVWIVTPKSGTSL 60

Qy 61 LOEVVYLVSQADPDEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQADPDEIGLNMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYMARNP KD LVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTMVQELARFLGV-----SCDKAQLAEALTEHCHQL 233
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTMVQELARFLGV-----SCDKAQLAEALTEHCHQL 233

Qy 234 VQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 237 -TASGPSSSLVCVSRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 286
```

```
RESULT 58
US-10-990-328-7685
; Sequence 7685, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7685
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328-7685

Query Match      82.8%; Score 1267; DB 39; Length 286;
Best Local Similarity 84.5%; Pred. No. 2.4e-126;
Matches 246; Conservative 7; Mismatches 26; Indels 12; Gaps 3;

QY 1 MAESAETPSTPGEFESKYFEHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
DB 1 MAESAETPSTPGEFESKYFEHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
QY 61 LOEVVYLSQGADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
DB 61 LOEVVYLSQGADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNPKDLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
DB 121 LHNGDSKVIYMARNPKDLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
QY 181 EFWEHRMDSNVFLKYEDMHRD-LVTWVEQLARFLGV-----SCDKAQLEALTEHCQOL 233
DB 181 EFWEHRMDSNVFLKYEDMHRD-LVTWVEQLARFLGV-----SCDKAQLEALTEHCQOL 233
QY 234 VDQCCNAEALPVGRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284
DB 234 VDQCCNAEALPVGRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284
QY 237 -TASGPSSSLVCVSRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 286
DB 237 -TASGPSSSLVCVSRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 286

RESULT 59
US-10-990-328A-7685
; Sequence 7685, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7685
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328A-7685

Query Match      82.8%; Score 1267; DB 39; Length 286;
Best Local Similarity 84.5%; Pred. No. 2.4e-126;
Matches 246; Conservative 7; Mismatches 26; Indels 12; Gaps 3;

QY 1 MAESAETPSTPGEFESKYFEHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
DB 1 MAESAETPSTPGEFESKYFEHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
QY 61 LOEVVYLSQGADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
DB 61 LOEVVYLSQGADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNPKDLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
DB 121 LHNGDSKVIYMARNPKDLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
QY 181 EFWEHRMDSNVFLKYEDMHRD-LVTWVEQLARFLGV-----SCDKAQLEALTEHCQOL 233
DB 181 EFWEHRMDSNVFLKYEDMHRD-LVTWVEQLARFLGV-----SCDKAQLEALTEHCQOL 233
QY 234 VDQCCNAEALPVGRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284
DB 234 VDQCCNAEALPVGRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284
QY 237 -TASGPSSSLVCVSRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 286
DB 237 -TASGPSSSLVCVSRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 286

RESULT 60
US-10-990-964-407
; Sequence 407, Application US/60499964
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN DRUG-METABOLIZING ENZYMES, METHODS OF
; FILE REFERENCE: CL001481
; CURRENT APPLICATION NUMBER: US/60/499,964
; CURRENT FILING DATE: 2003-09-04
; NUMBER OF SEQ ID NOS: 30431
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 407
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-964-407

Query Match      82.8%; Score 1267; DB 49; Length 286;
Best Local Similarity 84.5%; Pred. No. 2.4e-126;
Matches 246; Conservative 7; Mismatches 26; Indels 12; Gaps 3;

QY 1 MAESAETPSTPGEFESKYFEHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
DB 1 MAESAETPSTPGEFESKYFEHGVRLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
QY 61 LOEVVYLSQGADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
DB 61 LOEVVYLSQGADPDEIGLMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNPKDLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
DB 121 LHNGDSKVIYMARNPKDLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
QY 181 EFWEHRMDSNVFLKYEDMHRD-LVTWVEQLARFLGV-----SCDKAQLEALTEHCQOL 233
DB 181 EFWEHRMDSNVFLKYEDMHRD-LVTWVEQLARFLGV-----SCDKAQLEALTEHCQOL 233
QY 234 VDQCCNAEALPVGRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284
DB 234 VDQCCNAEALPVGRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 284
QY 237 -TASGPSSSLVCVSRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 286
DB 237 -TASGPSSSLVCVSRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTFFDYL 286

RESULT 61
PCT-US02-11654-172
; Sequence 172, Application PC/TUS0211654
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Farris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
```



```
US-11-154-298-172
; Sequence 172, Application US/11154298
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.01
; CURRENT APPLICATION NUMBER: US/11/154,298
; CURRENT FILING DATE: 2005-06-15
; PRIOR FILING DATE: 2002-04-10
; PRIOR FILING DATE: 2002-04-10
; PRIOR FILING DATE: 2002-04-10
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 172
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-154-298-172

Query Match      65.5%; Score 1002; DB 41; Length 218;
Best Local Similarity 96.4%; Pred. No. 4.5e-98;
Matches 186; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 92 POPGLDIIKELTSPRLIKSHLPYRPLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSLRT 151
Db 32 PSP-----ELTSPRLIKSHLPYRPLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSLRT 85

QY 152 MSYRTGTFQFCRRFMNDKLGYSWFHVEHQRMSNVLFKLYEDMHRDLVTWVEQLA 211
Db 86 MSYRTGTFQFCRRFMNDKLGYSWFHVEHQRMSNVLFKLYEDMHRDLVTWVEQLA 145

QY 212 RFLGVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 271
Db 146 RFLGVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 205

QY 272 KMGKCDLTFDFYL 284
Db 206 KMGKCDLTFDFYL 218

US-11-154-298-172

Query Match      65.5%; Score 1002; DB 41; Length 218;
Best Local Similarity 96.4%; Pred. No. 4.5e-98;
Matches 186; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 92 POPGLDIIKELTSPRLIKSHLPYRPLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSLRT 151
Db 32 PSP-----ELTSPRLIKSHLPYRPLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSLRT 85

QY 152 MSYRTGTFQFCRRFMNDKLGYSWFHVEHQRMSNVLFKLYEDMHRDLVTWVEQLA 211
Db 86 MSYRTGTFQFCRRFMNDKLGYSWFHVEHQRMSNVLFKLYEDMHRDLVTWVEQLA 145

QY 212 RFLGVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 271
Db 146 RFLGVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 205

QY 272 KMGKCDLTFDFYL 284
Db 206 KMGKCDLTFDFYL 218

US-11-155-906-172
; Sequence 172, Application US/11155906
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.02
; CURRENT APPLICATION NUMBER: US/11/155,906
; CURRENT FILING DATE: 2005-06-17
; PRIOR FILING DATE: 2002-04-10
; PRIOR FILING DATE: 2002-04-10
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 172
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-155-906-172

Query Match      65.5%; Score 1002; DB 41; Length 218;
Best Local Similarity 96.4%; Pred. No. 4.5e-98;
Matches 186; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 92 POPGLDIIKELTSPRLIKSHLPYRPLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSLRT 151
Db 32 PSP-----ELTSPRLIKSHLPYRPLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSLRT 85

QY 152 MSYRTGTFQFCRRFMNDKLGYSWFHVEHQRMSNVLFKLYEDMHRDLVTWVEQLA 211
Db 86 MSYRTGTFQFCRRFMNDKLGYSWFHVEHQRMSNVLFKLYEDMHRDLVTWVEQLA 145

QY 212 RFLGVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 271
Db 146 RFLGVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 205

QY 272 KMGKCDLTFDFYL 284
Db 206 KMGKCDLTFDFYL 218

US-11-156-231-172
; Sequence 172, Application US/11156231
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.03
; CURRENT APPLICATION NUMBER: US/11/156,231
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 172
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-156-231-172

Query Match      65.5%; Score 1002; DB 41; Length 218;
Best Local Similarity 96.4%; Pred. No. 4.5e-98;
Matches 186; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 92 POPGLDIIKELTSPRLIKSHLPYRPLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSLRT 151
Db 32 PSP-----ELTSPRLIKSHLPYRPLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSLRT 85

QY 152 MSYRTGTFQFCRRFMNDKLGYSWFHVEHQRMSNVLFKLYEDMHRDLVTWVEQLA 211
Db 32 PSP-----ELTSPRLIKSHLPYRPLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSLRT 85

QY 152 MSYRTGTFQFCRRFMNDKLGYSWFHVEHQRMSNVLFKLYEDMHRDLVTWVEQLA 211
```

Db 86 MSYGRGFECRRFNDKLGSGWFEHQEFWEHRMDSNVLFKYEDMHRDLVTWVEQIA 145  
Qy 212 RFLGVSCDKAQLAETHECHQLVDQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 271  
Db 146 RFLGVSCDKAQLAETHECHQLVDQCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVYKQ 205  
Qy 272 KMGKCDLTFDFYL 284  
Db 206 KMGKCDLTFDFYL 218

## RESULT 67

PCT-US02-11654-173  
; Sequence 173, Application PC/TUS0211654  
; GENERAL INFORMATION:  
; APPLICANT: Agensys, Inc.  
; APPLICANT: Jakobovits, Aya  
; APPLICANT: Challita-Bid, Pia M.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Morrison, Robert Kendall  
; APPLICANT: Raitano, Arthur B.  
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful  
; FILE REFERENCE: 51158-20040.40  
; CURRENT APPLICATION NUMBER: PCT/US02/11654  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: US 60/283,112  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US 60/282,739  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US 60/286,630  
; PRIOR FILING DATE: 2001-04-25  
; NUMBER OF SEQ ID NOS: 708  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 173  
; LENGTH: 171  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-11654-173

Query Match 56.7%; Score 867.5; DB 1; Length 171;  
Best Local Similarity 59.9%; Pred. No. 8.6e-84;  
Matches 170; Conservative 1; Mismatches 0; Indels 113; Gaps 1;  
Qy 1 MAESEAETPSTPGEFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60  
Db 1 MAESEAETPSTPGEFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 56  
Qy 61 LOEVVYLVSGADPDDEIGLNMNIDSQLVLEYVPPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Db 57 -----  
Qy 121 LHNGDSKVIYARNPKDLVVSYYQPHRSRLTMSYRGTFQECRRFMDKLGSGWFEHQV 180  
Db 57 -----VGYSWFEHQV 67  
Qy 181 EFWHRMDSNVLFKYEDMHRDLVTWVEQIAFLVSCDKAQLAETHECHQLVDQCCNA 240  
Db 68 EFWHRMDSNVLFKYEDMHRDLVTWVEQIAFLVSCDKAQLAETHECHQLVDQCCNA 127  
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284  
Db 128 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 171

## RESULT 68

US-10-121-024B-173  
; Sequence 173, Application US/10121024B  
; GENERAL INFORMATION:

; APPLICANT: Agensys, Inc.  
; APPLICANT: Jakobovits, Aya  
; APPLICANT: Challita-Bid, Pia M.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Morrison, Robert Kendall  
; APPLICANT: Raitano, Arthur B.  
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful  
; FILE REFERENCE: 51158-20040.00  
; CURRENT APPLICATION NUMBER: US/10/121,024B  
; CURRENT FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: US60/283,112  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US60/282,739  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US60/286,630  
; PRIOR FILING DATE: 2001-04-25  
; NUMBER OF SEQ ID NOS: 708  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 173  
; LENGTH: 171  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-121-024B-173  
Query Match 56.7%; Score 867.5; DB 31; Length 171;  
Best Local Similarity 59.9%; Pred. No. 8.6e-84;  
Matches 170; Conservative 1; Mismatches 0; Indels 113; Gaps 1;  
Qy 1 MAESEAETPSTPGEFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60  
Db 1 MAESEAETPSTPGEFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 56  
Qy 61 LOEVVYLVSGADPDDEIGLNMNIDSQLVLEYVPPQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Db 57 -----  
Qy 121 LHNGDSKVIYARNPKDLVVSYYQPHRSRLTMSYRGTFQECRRFMDKLGSGWFEHQV 180  
Db 57 -----VGYSWFEHQV 67  
Qy 181 EFWHRMDSNVLFKYEDMHRDLVTWVEQIAFLVSCDKAQLAETHECHQLVDQCCNA 240  
Db 68 EFWHRMDSNVLFKYEDMHRDLVTWVEQIAFLVSCDKAQLAETHECHQLVDQCCNA 127  
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284  
Db 128 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 171  
RESULT 69  
US-10-121-024C-173  
; Sequence 173, Application US/10121024C  
; GENERAL INFORMATION:  
; APPLICANT: Jakobovits, Aya  
; APPLICANT: Challita-Bid, Pia M.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Morrison, Robert Kendall  
; APPLICANT: Raitano, Arthur B.  
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Useful  
; FILE REFERENCE: 51158-20040.00  
; CURRENT APPLICATION NUMBER: US/10/121,024C  
; CURRENT FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: US60/283,112  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US60/282,739



```

; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 173
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-024C-173

Query Match          56.7%; Score 867.5; DB 31; Length 171;
Best Local Similarity 59.9%; Pred. No. 8.6e-84;
Matches 170; Conservative 1; Mismatches 0; Indels 113; Gaps 1;

Qy 1 MAESEAETSTPGEFESKYPFEGHGVRLPPFCRGKMBEIANFPVRPSDWMIVTPKSGTSL 60
Db 1 MAESEAETSTPGEFESKYPFEGHGVRLPPFCRGKMBEIANFPVRPSDWMIVTPKS---- 56
;

Qy 61 LQEVVYLVSGADPDDEIGLNMIDEQLPVLEYPOPGLDIIKELTSPLRIKSHLPYRFLPSD 120
Db 57 -----
;

Qy 121 LHNGDSKVIYARNPKDLVSVYYQFHRSRLTMSYRGTFQFCRRFNMNDKLGYSWFEHVQ 180
Db 57 -----
;

Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 240
Db 68 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 127
;

Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db 128 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 171
;

RESULT 70
US-10-573-989-18
; Sequence 18, Application US/10573989
; GENERAL INFORMATION:
; APPLICANT: Evotec NeuroSciences GmbH
; APPLICANT: Von Der Kammer, Heinz
; APPLICANT: Pohnler, Johannes
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC USE OF A SULFOTRANSFERASE
; FILE REFERENCE: 2335, 0140000/SRL/KPO
; CURRENT APPLICATION NUMBER: US/10/573,989
; CURRENT FILING DATE: 2006-03-30
; PRIOR APPLICATION NUMBER: PCT/EP2004/052353
; PRIOR FILING DATE: 2004-09-29
; PRIOR APPLICATION NUMBER: 60/506,775
; PRIOR FILING DATE: 2003-09-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-573-989-18

Query Match          56.7%; Score 867.5; DB 35; Length 171;
Best Local Similarity 59.9%; Pred. No. 8.6e-84;
Matches 170; Conservative 1; Mismatches 0; Indels 113; Gaps 1;

Qy 1 MAESEAETSTPGEFESKYPFEGHGVRLPPFCRGKMBEIANFPVRPSDWMIVTPKSGTSL 60
Db 1 MAESEAETSTPGEFESKYPFEGHGVRLPPFCRGKMBEIANFPVRPSDWMIVTPKS---- 56
;

Qy 61 LQEVVYLVSGADPDDEIGLNMIDEQLPVLEYPOPGLDIIKELTSPLRIKSHLPYRFLPSD 120
Db 57 -----
;

Qy 121 LHNGDSKVIYARNPKDLVSVYYQFHRSRLTMSYRGTFQFCRRFNMNDKLGYSWFEHVQ 180
Db 57 -----
;

; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 173
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-154-298-173

Query Match          56.7%; Score 867.5; DB 41; Length 171;
Best Local Similarity 59.9%; Pred. No. 8.6e-84;
Matches 170; Conservative 1; Mismatches 0; Indels 113; Gaps 1

Qy 1 MAESEAETSTPGEFESKYPFEGHGVRLPPFCRGKMBEIANFPVRPSDWMIVTPKSGTSL 60
Db 1 MAESEAETSTPGEFESKYPFEGHGVRLPPFCRGKMBEIANFPVRPSDWMIVTPKS---- 56
;

Qy 61 LQEVVYLVSGADPDDEIGLNMIDEQLPVLEYPOPGLDIIKELTSPLRIKSHLPYRFLPSD 120
Db 57 -----
;

Qy 121 LHNGDSKVIYARNPKDLVSVYYQFHRSRLTMSYRGTFQFCRRFNMNDKLGYSWFEHVQ 180
Db 57 -----
;

Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 240
Db 68 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 127
;

Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db 128 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 171
;

RESULT 71
US-11-154-298-173
; Sequence 173, Application US/11154298
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.01
; CURRENT APPLICATION NUMBER: US/11/154,298
; CURRENT FILING DATE: 2005-06-15
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 173
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-154-298-173

Query Match          56.7%; Score 867.5; DB 41; Length 171;
Best Local Similarity 59.9%; Pred. No. 8.6e-84;
Matches 170; Conservative 1; Mismatches 0; Indels 113; Gaps 1

Qy 1 MAESEAETSTPGEFESKYPFEGHGVRLPPFCRGKMBEIANFPVRPSDWMIVTPKSGTSL 60
Db 1 MAESEAETSTPGEFESKYPFEGHGVRLPPFCRGKMBEIANFPVRPSDWMIVTPKS---- 56
;

Qy 61 LQEVVYLVSGADPDDEIGLNMIDEQLPVLEYPOPGLDIIKELTSPLRIKSHLPYRFLPSD 120
Db 57 -----
;

Qy 121 LHNGDSKVIYARNPKDLVSVYYQFHRSRLTMSYRGTFQFCRRFNMNDKLGYSWFEHVQ 180
Db 57 -----
;

Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 240
Db 68 EFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNA 127
;

Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db 128 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 171
;

RESULT 72
US-11-155-906-173
; Sequence 173, Application US/11155906
; GENERAL INFORMATION:

```

```

; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.02
; CURRENT APPLICATION NUMBER: US/11/155,906
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 173
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-155-906-173

Query Match          56.7%; Score 867.5; DB 41; Length 171;
Best Local Similarity 59.9%; Pred. No. 8.6e-84;
Matches 170; Conservative 1; Mismatches 0; Indels 113; Gaps 1

QY 1 MAESEAETPTPGFESKYFFHGVRLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60
DB 1 MAESEAETPTPGFESKYFFHGVRLPPFCRGKMEETANFVRPSDVWIVTPKS----- 56
QY 61 LQEVVYLVSQGADPEIGLANNIDBQLPVLEYQPQGLDIIKELTSPLIKSHLPYRFLPSD 120
DB 57 ----- 56
QY 121 LHNGDSKVIYARNPKDLVSYQFHRSLRTWSYRGTFQFCRRPMNDKLGVSWFHVQ 180
DB 57 -----VGYGSWFHVQ 67
QY 181 EFWEHRMDSNVLFLKYEDMHDRDLVTWVEQLARFLGVSCDKAQLEALTEHCHQLVDQCNA 240
DB 68 EFWEHRMDSNVLFLKYEDMHDRDLVTWVEQLARFLGVSCDKAQLEALTEHCHQLVDQCNA 127
QY 241 EALPVGRGRVGLWKDIFTVSNNEKFDLVYKQKMGKCDLTFDFYL 284
DB 128 EALPVGRGRVGLWKDIFTVSNNEKFDLVYKQKMGKCDLTFDFYL 171

RESULT 73
US-11-156-231-173
; Sequence 173, Application US/11156231
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Ge, Wangmao
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Morrison, Robert Kendall
; APPLICANT: Raitano, Arthur B.
; TITLE OF INVENTION: METHODS OF INDUCING AN IMMUNE RESPONSE
; FILE REFERENCE: 51158-20040.03
; CURRENT APPLICATION NUMBER: US/11/156,231
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US60/283,112
; PRIOR FILING DATE: 2001-04-10

```

RESULT 75  
US-10-170-205E-27771  
; Sequence 27771, Application US/10170205E  
; GENERAL INFORMATION:  
; APPLICANT: ADAMS, Mark  
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN  
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF  
; FILE REFERENCE: CLO01381  
; CURRENT APPLICATION NUMBER: US/10/170,205E  
; CURRENT FILING DATE: 2002-06-13  
; NUMBER OF SEQ ID NOS: 40312  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 27771  
; LENGTH: 194  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-170-205E-27771

Query Match 52.5%; Score 804; DB 31; Length 194;  
Best Local Similarity 91.0%; Pred. No. 6.7e-77;  
Matches 151; Conservative 1; Mismatches 8; Indels 6; Gaps 1;

Qy 92 POPGLDIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMAENPKDLVVSYYQFHRSLRT 151  
Db 32 PSP-----ELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMAENPKDLVVSYYQFHRSLRT 85

Qy 152 MSYRGTFQFCRRFWMNDKLGYSWFHEHVOEFWEHRMDSNVLFKLYEDMHRDLVTMVEOLA 211  
Db 86 MSYRGTFQFCRRFWMNDKLGYSWFHEHVOEFWEHRMDSNVLFKLYEDMHRDLVTMVEOLA 145

Qy 212 RFLGVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRGRVGLWKDIF 257  
Db 146 RFLGVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRAHCVFARKIF 191

RESULT 76  
US-09-791-537-24613  
; Sequence 24613, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 24613  
; LENGTH: 298  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-24613

Query Match 32.4%; Score 495; DB 27; Length 298;  
Best Local Similarity 35.5%; Pred. No. 1.8e-43;  
Matches 104; Conservative 54; Mismatches 117; Indels 18; Gaps 3;

Qy 6 AETPSTPGEFESKYFEFHGRVLRPPFCRGKMBEIANFPVRPSDVWITVTPKSGTSLQEVV 65  
Db 2 AQNPNSMEPLRKLPLVPVKGIPLIKYFAETMEQLQNFATWPDVDLISTYPKSGTWNWSEIM 61

Qy 66 YLVSGADPDEIGLMNIDEQLPVLEYP-----QPLGDIKELTSPRLIKSHLPYRFLPSDL 121  
Db 62 DMIYQGGKLDKGRAPVYARIPFLFSCPGVPVPGLETTLKETPAPRIIKTHLPLSLLOSL 121

Qy 122 HNGDSKVIYMAENPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFWMNDKLGYSWFHEHVOE 181  
Db 122 LDQIKVIYARNADVVVSYYNFYKMAKLHPDPTGWESFLENFMDGKVSYSWYQHVKE 181

Qy 182 FWEHRMDSNVLFKLYEDMHRDLVTMVEQLARFLGVSCDKAQLEALTEHC---HQLVDQCC 238  
Db 182 WWELRRTHPVLVLYFYEDMKENPKREIKKILFELGRSLPEETVDLIVVHTSFKKMKENPMA 241

Qy 239 NAEALP-----VGRGRVGLWKDIFTVSMNEKFDLYVKQKMGKCDLTF 280  
Db 242 NVTTIPTVMDHTIYPPFMRKGTIGDWKNTFTVAQSEHFDHAHYAKLMTGCDFTF 294

RESULT 78  
US-09-791-537-66629  
; Sequence 66629, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 66629  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-66629

Query Match 32.4%; Score 495; DB 27; Length 298;  
Best Local Similarity 35.5%; Pred. No. 1.8e-43;  
Matches 104; Conservative 54; Mismatches 117; Indels 18; Gaps 3;

Qy 6 AETPSTPGEFESKYFEFHGRVLRPPFCRGKMBEIANFPVRPSDVWITVTPKSGTSLQEVV 65  
Db 2 AQNPNSMEPLRKLPLVPVKGIPLIKYFAETMEQLQNFATWPDVDLISTYPKSGTWNWSEIM 61

Qy 66 YLVSGADPDEIGLMNIDEQLPVLEYP-----QPLGDIKELTSPRLIKSHLPYRFLPSDL 121  
Db 62 DMIYQGGKLDKGRAPVYARIPFLFSCPGVPVPGLETTLKETPAPRIIKTHLPLSLLOSL 121

Qy 122 HNGDSKVIYMAENPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFWMNDKLGYSWFHEHVOE 181  
Db 122 LDQIKVIYARNADVVVSYYNFYKMAKLHPDPTGWESFLENFMDGKVSYSWYQHVKE 181

Qy 182 FWEHRMDSNVLFKLYEDMHRDLVTMVEQLARFLGVSCDKAQLEALTEHC---HQLVDQCC 238

Db 182 WWELRRTHPVLVLYFYEDMKENPKREIKKILFELGRSLPEETVDLIVVHTSFKKMKENPMA 241

Qy 239 NAEALP-----VGRGRVGLWKDIFTVSMNEKFDLYVKQKMGKCDLTF 280

Db 242 NVTTIPTVMDHTIYPPFMRKGTIGDWKNTFTVAQSEHFDHAHYAKLMTGCDFTF 294

RESULT 77  
US-09-791-537-26347  
; Sequence 26347, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 26347  
; LENGTH: 298  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-26347

Query Match 32.4%; Score 495; DB 27; Length 298;  
Best Local Similarity 35.5%; Pred. No. 1.8e-43;  
Matches 104; Conservative 54; Mismatches 117; Indels 18; Gaps 3;

Qy 6 AETPSTPGEFESKYFEFHGRVLRPPFCRGKMBEIANFPVRPSDVWITVTPKSGTSLQEVV 65  
Db 2 AQNPNSMEPLRKLPLVPVKGIPLIKYFAETMEQLQNFATWPDVDLISTYPKSGTWNWSEIM 61

Qy 66 YLVSGADPDEIGLMNIDEQLPVLEYP-----QPLGDIKELTSPRLIKSHLPYRFLPSDL 121  
Db 62 DMIYQGGKLDKGRAPVYARIPFLFSCPGVPVPGLETTLKETPAPRIIKTHLPLSLLOSL 121

Qy 122 HNGDSKVIYMAENPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFWMNDKLGYSWFHEHVOE 181  
Db 122 LDQIKVIYARNADVVVSYYNFYKMAKLHPDPTGWESFLENFMDGKVSYSWYQHVKE 181

Qy 182 FWEHRMDSNVLFKLYEDMHRDLVTMVEQLARFLGVSCDKAQLEALTEHC---HQLVDQCC 238  
Db 182 WWELRRTHPVLVLYFYEDMKENPKREIKKILFELGRSLPEETVDLIVVHTSFKKMKENPMA 241

Qy 239 NAEALP-----VGRGRVGLWKDIFTVSMNEKFDLYVKQKMGKCDLTF 280  
Db 242 NVTTIPTVMDHTIYPPFMRKGTIGDWKNTFTVAQSEHFDHAHYAKLMTGCDFTF 294

RESULT 78  
US-09-791-537-66629  
; Sequence 66629, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 66629  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-66629



Db 445 AHPEGTWDSFLEKFMAGEVSYGSHVQVQEWELSRTHPVLYLFYEDMKENPKREIQKI 504  
Qy 211 ARFLGVSCDKAQLEALTEHC-----HQLVQCCNABALP-VGRGRVGL 252  
Db 505 LEFVGRSLPEETMDPMVQHTSFKEKKNPMNTYTTVPQELMDHSIS-----PFMRKGMAGD 560  
Qy 253 WKDIFTVMNEKFDLVYKQKMGKCDLTFDFYL 284  
Db 561 WKTFTVAQNERFDADYAEKMGAGCSLSRPFEL 592

RESULT 81  
US-60-229-512-201  
; Sequence 201, Application US/60229512  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN PHASE II  
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, NUCLEIC ACID MOLECULES ENCODING  
; TITLE OF INVENTION: HUMAN DRUG-METABOLIZING PHASE II PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL000770  
; CURRENT APPLICATION NUMBER: US/60/229,512  
; CURRENT FILING DATE: 2000-09-05  
; NUMBER OF SEQ ID NOS: 516  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 201  
; LENGTH: 592  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-229-512-201

Query Match 31.4%; Score 480; DB 47; Length 592;  
Best Local Similarity 33.7%; Pred. No. 2e-41;  
Matches 112; Conservative 54; Mismatches 112; Indels 54; Gaps 8;  
Qy 3 ESEATPSTP-----GEFSKYFEPHGVL-----PP--FCRG-----K 34  
Db 265 BEGREQPTPTTPTLSPCROGENSGRELRNMELIQTSPPLLEYVKGVPPLIKYFAEA 324  
Qy 35 MEEIANFPVRPSDVWIVTPKSGTSLLOEVVYVLSQGDADPDDEIGLNMNIDEQLPVLEYQP 94  
Db 325 LGPQSFQARDDLLINTYPSKGTWVSQILDMYQGGDLKCNRAPYVRVPFLEVNDP 384  
Qy 95 ----GLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYMARNPOLDVSYQPHRSLR 150  
Db 385 GEPSGLETLDKTPPPRLIKSHLPLALLPQTLDDQKVKVYVARNPKDVAVSYYHPRMEK 444  
Qy 151 TMSYRGTFQECRRFMNDKLGYSWFHVFQEFWEHRMDSNVLFKYEDMHRDLVTMVQQL 210  
Db 445 AHPEGTWDSFLEKFMAGEVSYGSHVQVQEWELSRTHPVLYLFYEDMKENPKREIQKI 504  
Qy 211 ARFLGVSCDKAQLEALTEHC-----HQLVQCCNABALP-VGRGRVGL 252  
Db 505 LEFVGRSLPEETMDPMVQHTSFKEKKNPMNTYTTVPQELMDHSIS-----PFMRKGMAGD 560  
Qy 253 WKDIFTVMNEKFDLVYKQKMGKCDLTFDFYL 284  
Db 561 WKTFTVAQNERFDADYAEKMGAGCSLSRPFEL 592

RESULT 82  
US-09-791-537-52308  
; Sequence 52308, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 52308  
; LENGTH: 295  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-791-537-52308  
Query Match 31.3%; Score 479; DB 27; Length 295;  
Best Local Similarity 36.6%; Pred. No. 9.2e-42;  
Matches 102; Conservative 50; Mismatches 101; Indels 26; Gaps 4;  
Qy 24 GVRLPFFCRGMEETANFPVRPSDVWIVTPKSGTSLLOEVVYVLSQGDADPDDEIGLNMNID 83  
Db 17 GVPLIKYFAEALGPLQSFQARPDLLINTYPSKGTWVSQILDMYQGGDLKCNRAPYI 76  
Qy 84 EQLPVLEYQP-----GLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYMARNPOLDV 139  
Db 77 VRVPFLEVNDPGEPSGLETLDKTPPPRLIKSHLPLALLPQTLDDQKVKVYVARNPKDVA 136  
Qy 140 VSYQPHRSLRMTSYRGTFQECRRFMNDKLGYSWFHVFQEFWEHRMDSNVLFKYEDM 199  
Db 137 VSYTHFRMEKAHPPEPGTWDSEFLEKFMAGEVSYGSHVQVQEWELSRTHPVLYLFYEDM 196  
Qy 200 HRDLVTMVQELARFLGVSCDKAQLEALTEHC-----HQLVQCCNABEA 242  
Db 197 KENPKREIQKILEFVGRSLPBEETVDLMVVEHTSFKEKMKTPMTNTYTVRRELMDSIS---- 253  
Qy 243 LP-VGRGRVGLWKDIFTVMNEKFDLVYKQKMGKCDLTF 280  
Db 254 -PFMRKGMAGDKWKTFTTVAQNERFDADYAEKMGAGCSLTF 291

RESULT 83  
PCT-US01-08631-57913  
; Sequence 57913, Application PC/TUS0108631  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 21272-049  
; CURRENT APPLICATION NUMBER: PCT/US01/08631  
; CURRENT FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 57913  
; LENGTH: 1305  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: DOMAIN  
; LOCATION: (649)..(695)  
; OTHER INFORMATION: Serine/threonine specific protein phosphatases proteins  
; OTHER INFORMATION: domain identified by eMATRIX, accession number BL00125B, p-value=  
; OTHER INFORMATION: 1.000e-40, raw score of 21.48  
; NAME/KEY: DOMAIN  
; LOCATION: (593)..(850)  
; OTHER INFORMATION: Ser/Thr protein phosphatase domain identified by Pfam,  
; OTHER INFORMATION: accession name Stposphatase, E-value=2.6e-196, Pfam score of 665  
PCT-US01-08631-57913

Query Match 31.0%; Score 474; DB 1; Length 1305;  
Best Local Similarity 33.5%; Pred. No. 2.7e-40;  
Matches 110; Conservative 54; Mismatches 110; Indels 54; Gaps 8;  
Qy 3 ESEATPSTP-----GEFSKYFEPHGVL-----PP--FCRG-----K 34  
Db 265 BEGREQPTPTTPTLSPCROGENSGRELRNMELIQTSPPLLEYVKGVPPLIKYFAEA 324  
Qy 35 MEEIANFPVRPSDVWIVTPKSGTSLLOEVVYVLSQGDADPDDEIGLNMNIDEQLPVLEYQP 94

Db 325 LGPLOSFOARPDLLINTPKSGTTWVSQLDLMIVQGGDLKCNRPAPYVVRVFFLEVNDP 384  
Qy 95 ---GLDIIKELTSPRLKSHLPVRFPSDLHNGDSKVIYMARNPDKLVSYQPHRSR 150  
Db 385 GEPSGLETLKOTPPRLIKSHLPLALLPQTLLDQKVYVYARNPKDVAVSYYHFRMEK 444  
Qy 151 TMSYRGTFQECRRRMNDKLGVSFWFHVQFWEHRMDSNVLFKLYEDMDHRDLVTMVQOL 210  
Db 445 AHPEPGTWDSELEKFMAGEVSGSWYQHVQEWELSRTHPVLVLYFYEDMKENPKREIQKI 504  
Qy 211 ARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAEALP-VGRGRVGL 252  
Db 505 LEFVGRSLPEETMDPMVQHTSFKETKNPMNTYTTVPQELMDHSIS-----PFMRKGMAGD 560  
Qy 253 WKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
Db 561 WKTFTVAQNERFDADYAEKMGAGCSLSF 588  
RESULT 84  
US-10-450-763-57913  
; Sequence 57913, Application US/10450763  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Cuetom  
; SEQ ID NO 57913  
; LENGTH: 1305  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: DOMAIN  
; LOCATION: (649)..(695)  
; OTHER INFORMATION: Serine/threonine specific protein phosphatases proteins  
; OTHER INFORMATION: domain identified by eMATRIX, accession number BL001258, p-value=  
; OTHER INFORMATION: 1.000e-40, raw score of 21.48  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (593)..(850)  
; OTHER INFORMATION: Ser/Thr protein phosphatase domain identified by Pfam,  
; OTHER INFORMATION: accession name Stphosphatase, E-value=2.6e-196, Pfam score of 665  
US-10-450-763-57913

Query Match 31.0%; Score 474; DB 34; Length 1305;  
Best Local Similarity 33.5%; Pred. No. 2.7e-40;  
Matches 110; Conservative 54; Mismatches 110; Indels 54; Gaps 8;  
Qy 3 ESEATPTTP-----GEFSKYFEFHGVRL-----PP--FCRG-----K 34  
Db 265 EEGREQPTPTTILSPCRQGENSGSRELNMELIQTSPRPLEYVKGVP LIKYFAEA 324  
Qy 35 MEEIANFPVRSDWIVTPKSGTSLLOEVVYVLSQGDPPDEIGLNMIDQOLPVLEYQRP 94  
Db 325 LGPLOSFOARPDLLINTPKSGTTWVSQLDLMIVQGGDLKCNRPAPYVVRVFFLEVNDP 384  
Qy 95 ---GLDIIKELTSPRLKSHLPVRFPSDLHNGDSKVIYMARNPDKLVSYQPHRSR 150  
Db 385 GEPSGLETLKOTPPRLIKSHLPLALLPQTLLDQKVYVYARNPKDVAVSYYHFRMEK 444  
Qy 151 TMSYRGTFQECRRRMNDKLGVSFWFHVQFWEHRMDSNVLFKLYEDMDHRDLVTMVQOL 210  
Db 445 AHPEPGTWDSELEKFMAGEVSGSWYQHVQEWELSRTHPVLVLYFYEDMKENPKREIQKI 504

Qy 211 ARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAEALP-VGRGRVGL 252  
Db 505 LEFVGRSLPEETMDPMVQHTSFKETKNPMNTYTTVPQELMDHSIS-----PFMRKGMAGD 560  
Qy 253 WKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
Db 561 WKTFTVAQNERFDADYAEKMGAGCSLSF 588  
RESULT 85  
PCT-US01-41986-17  
; Sequence 17, Application PC/TUS0141986  
; GENERAL INFORMATION:  
; APPLICANT: Trustees of Boston University  
; TITLE OF INVENTION: A NEURAL SPECIFIC CYTOSOLIC  
; FILE REFERENCE: SULFOTRANSFERASE FOR DRUG SCREENING  
; CURRENT APPLICATION NUMBER: PCT/US01/41986  
; CURRENT FILING DATE: 2001-09-04  
; PRIOR APPLICATION NUMBER: 60/229,929  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 295  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: DOMAIN  
; LOCATION: (472)..(52e-41)  
; OTHER INFORMATION: 30.8%; Score 472; DB 1; Length 295;  
; OTHER INFORMATION: Best Local Similarity 35.8%; Pred. No. 5.2e-41;  
; OTHER INFORMATION: Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;  
Qy 24 GVRLLPPFCRGKMEIANPFPVRPSDWMIVTPKSGTSLLOEVVYVLSQGDPPDEIGLNMID 83  
Db 17 GVPLIKYPAEALGPLQSFQARPDLLINTPKSGTTWVSQLDLMIVQGGDLKCNRPAPY 76  
Qy 84 EQLPVLEYQPP---GLDIIKELTSPRLKSHLPVRFPSDLHNGDSKVIYMARNPDKLV 139  
Db 77 VRVPFLEVNDPGEPSGLETTLKTPPRLIKSHLPLALLPQTLLDQKVYVYARNPKDVA 136  
Qy 140 VSYTFQHRSLRTMSYRGTFQECRRFMNDKLGVSFWFHVQFWEHRMDSNVLFKLYEDM 199  
Db 137 VSYVHFHMEKHAPEPGTWDSELEKFMAGEVSGSWYQHVQEWELSRTHPVLVLYFYEDM 196  
Qy 200 HRDLVTMVQELARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAE 242  
Db 197 KENPKREIQKILEFVGRSLPEETMDPMVQHTSFKEMKKNPMNTYTTVPQELMDHSIS--- 253  
Qy 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
Db 254 -PFMRKGMAGDWKTTFTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 86  
PCT-US05-01474-4  
; Sequence 4, Application PC/TUS0501474  
; GENERAL INFORMATION:  
; APPLICANT: Squires, James E.  
; APPLICANT: Lin, Zhihong E.  
; TITLE OF INVENTION: Genetic Markers for Skatole Metabolism  
; FILE REFERENCE: P06153U03  
; CURRENT APPLICATION NUMBER: PCT/US05/01474  
; CURRENT FILING DATE: 2005-01-18  
; PRIOR APPLICATION NUMBER: 10/769,507  
; PRIOR FILING DATE: 2004-01-30  
; PRIOR APPLICATION NUMBER: 10/024,628  
; PRIOR FILING DATE: 2001-11-23  
; PRIOR APPLICATION NUMBER: 09/288,037  
; PRIOR FILING DATE: 1999-04-08  
; PRIOR APPLICATION NUMBER: 60/081,037  
; PRIOR FILING DATE: 1998-04-08  
; NUMBER OF SEQ ID NOS: 14

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-01474-4

Query Match      30.8%; Score 472; DB 1; Length 295;
Best Local Similarity 35.8%; Pred. No. 5.2e-41;
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLLPFCRGMEEIANPVRPSDVWIVTPKSGTSLLOEVVYLVSQGADPDEIGLMNID 83
DB 17 GVLPIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDMIYQGGDLEKCNRAPIY 76

QY 84 EQLPVLEYPOP---GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNP KD LV 139
DB 77 VRVPFLEVNDGPEPSGLETLDKTPPRLIKSHLPLALLPQTLDDQKVKVYVARNPKDVA 136

QY 140 VSYIQFHRSRLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDSNVLFKYEDM 199
DB 137 VSYTHFRMEKAHPPEPGTWSDFLEKFMAGEVSYGSYQHWQEWELSRTHPVLYLFYEDM 196

QY 200 HRDLVTWVEQLARFLGVSCDKAQLAALTEHC-----HQLVDQCCNAEA 242
DB 197 KENPKREIQKILEFVGRSLPEETDMFVQHTSFKEMKKNPMTNVTYTPQELMDHSIS--- 253

QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
DB 254 -PFMRKGMAGDWKTTFTVAQNERFDADYAEKMGACCSLSF 291

RESULT 87
US-08-484-878-7
; Sequence 7, Application US/08484878
; GENERAL INFORMATION:
; APPLICANT: Beer-Romero, Peggy
; APPLICANT: Draetta, Giulio
; APPLICANT: Rolfe, Mark
; TITLE OF INVENTION: E6AP-Binding Proteins
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,878
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-044
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 295 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-484-878-7

Query Match      30.8%; Score 472; DB 14; Length 295;
Best Local Similarity 35.8%; Pred. No. 5.2e-41;
```

```
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLLPFCRGMEEIANPVRPSDVWIVTPKSGTSLLOEVVYLVSQGADPDEIGLMNID 83
DB 17 GVLPIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDMIYQGGDLEKCNRAPIY 76

QY 84 EQLPVLEYPOP---GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNP KD LV 139
DB 77 VRVPFLEVNDGPEPSGLETLDKTPPRLIKSHLPLALLPQTLDDQKVKVYVARNPKDVA 136

QY 140 VSYIQFHRSRLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDSNVLFKYEDM 199
DB 137 VSYTHFRMEKAHPPEPGTWSDFLEKFMAGEVSYGSYQHWQEWELSRTHPVLYLFYEDM 196

QY 200 HRDLVTWVEQLARFLGVSCDKAQLAALTEHC-----HQLVDQCCNAEA 242
DB 197 KENPKREIQKILEFVGRSLPEETDMFVQHTSFKEMKKNPMTNVTYTPQELMDHSIS--- 253

QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
DB 254 -PFMRKGMAGDWKTTFTVAQNERFDADYAEKMGACCSLSF 291

RESULT 88
US-09-791-537-67010
; Sequence 67010, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 67010
; LENGTH: 295
; TYPE: PRT
; ORGANISM: pdb 1CJWA
US-09-791-537-67010

Query Match      30.8%; Score 472; DB 27; Length 295;
Best Local Similarity 35.8%; Pred. No. 5.2e-41;
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLLPFCRGMEEIANPVRPSDVWIVTPKSGTSLLOEVVYLVSQGADPDEIGLMNID 83
DB 17 GVLPIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDMIYQGGDLEKCNRAPIY 76

QY 84 EQLPVLEYPOP---GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNP KD LV 139
DB 77 VRVPFLEVNDGPEPSGLETLDKTPPRLIKSHLPLALLPQTLDDQKVKVYVARNPKDVA 136

QY 140 VSYIQFHRSRLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDSNVLFKYEDM 199
DB 137 VSYTHFRMEKAHPPEPGTWSDFLEKFMAGEVSYGSYQHWQEWELSRTHPVLYLFYEDM 196

QY 200 HRDLVTWVEQLARFLGVSCDKAQLAALTEHC-----HQLVDQCCNAEA 242
DB 197 KENPKREIQKILEFVGRSLPEETDMFVQHTSFKEMKKNPMTNVTYTPQELMDHSIS--- 253

QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
DB 254 -PFMRKGMAGDWKTTFTVAQNERFDADYAEKMGACCSLSF 291

RESULT 89
US-10-170-205E-29903
; Sequence 29903, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
```







```
Qy 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVSQADPDEIGLNMID 199
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 137 VSYHFRHMEKAHPPEPGTWSDFLEKFMAGEVSYGSHYQVQEWELSTHVPVLYLFYEDM 196
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 200 HRDLVTMVEQLARFLGVSCDKAQLBALTEHC-----HQLVDQCCNAB 242
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 197 KENPKREIQKILEFVGRSLPEETMDFMVQHTSFKEKMKKNPMTNVTVPQELMDHSIS--- 253
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 254 -PFMRKGMAGDKWKTFTTVAQNERFDADYAEKMGAGCSLSF 291
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

## RESULT 96

US-10-769-507-4

```
; Sequence 4, Application US/10769507
; GENERAL INFORMATION:
; APPLICANT: Squires, James E.
; TITLE OF INVENTION: Genetic Markers for Skatole Metabolism
; FILE REFERENCE: P06153US03
; CURRENT APPLICATION NUMBER: US/10/769,507
; PRIOR FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: 10/024,628
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: 09/288,037
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/081,037
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-769-507-4
```

```
Query Match 30.8%; Score 472; DB 37; Length 295;
Best Local Similarity 35.8%; Pred. No. 5.2e-41;
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;

Qy 24 GVRLLPPFCRGKMEIANTFVRPSDVWIVTPYKSGTSLLOEVVYVSQADPDEIGLNMID 83
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 17 GVPLIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDIMYQGGDLEKCNRAPIY 76
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 84 EQLPVLEYPOP-----GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNPDLV 139
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 77 VRVPFLEVNDPGEPSGLETLDTPPPRLIKSHLPLALLPOTLLDQKVKVYVARNPKDVA 136
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVSQADPDEIGLNMID 199
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 137 VSYHFRHMEKAHPPEPGTWSDFLEKFMAGEVSYGSHYQVQEWELSTHVPVLYLFYEDM 196
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 200 HRDLVTMVEQLARFLGVSCDKAQLBALTEHC-----HQLVDQCCNAB 242
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 197 KENPKREIQKILEFVGRSLPEETMDFMVQHTSFKEKMKKNPMTNVTVPQELMDHSIS--- 253
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 254 -PFMRKGMAGDKWKTFTTVAQNERFDADYAEKMGAGCSLSF 291
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

## RESULT 97

```
US-11-033-030-21
; Sequence 21, Application US/11033030
; GENERAL INFORMATION:
; APPLICANT: SMITH, ROBERT
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
; FILE REFERENCE: PHA-007.01
; CURRENT APPLICATION NUMBER: US/11/033,030
; CURRENT FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: US/09/854,122
```

```
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/202,529
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 21
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-033-030-21

Query Match 30.8%; Score 472; DB 40; Length 295;
Best Local Similarity 35.8%; Pred. No. 5.2e-41;
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;

Qy 24 GVRLLPPFCRGKMEIANTFVRPSDVWIVTPYKSGTSLLOEVVYVSQADPDEIGLNMID 83
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 17 GVPLIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDIMYQGGDLEKCNRAPIY 76
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 84 EQLPVLEYPOP-----GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNPDLV 139
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 77 VRVPFLEVNDPGEPSGLETLDTPPPRLIKSHLPLALLPOTLLDQKVKVYVARNPKDVA 136
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVSQADPDEIGLNMID 199
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 137 VSYHFRHMEKAHPPEPGTWSDFLEKFMAGEVSYGSHYQVQEWELSTHVPVLYLFYEDM 196
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 200 HRDLVTMVEQLARFLGVSCDKAQLBALTEHC-----HQLVDQCCNAB 242
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 197 KENPKREIQKILEFVGRSLPEETMDFMVQHTSFKEKMKKNPMTNVTVPQELMDHSIS--- 253
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 254 -PFMRKGMAGDKWKTFTTVAQNERFDADYAEKMGAGCSLSF 291
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

## RESULT 98

```
US-60-568-073-1205
; Sequence 1205, Application US/60568073
; GENERAL INFORMATION:
; APPLICANT: Huang, Fei
; APPLICANT: Han, Xia
; APPLICANT: Shaw, Peter
; APPLICANT: Clark, Edwin
; TITLE OF INVENTION: COLON CANCER BIOMARKERS
; FILE REFERENCE: 10205 PSP
; CURRENT APPLICATION NUMBER: US/60/568,073
; CURRENT FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 1305
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1205
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-568-073-1205

Query Match 30.8%; Score 472; DB 50; Length 295;
Best Local Similarity 35.8%; Pred. No. 5.2e-41;
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;

Qy 24 GVRLLPPFCRGKMEIANTFVRPSDVWIVTPYKSGTSLLOEVVYVSQADPDEIGLNMID 83
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 17 GVPLIKYFAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDIMYQGGDLEKCNRAPIY 76
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 84 EQLPVLEYPOP-----GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNPDLV 139
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 77 VRVPFLEVNDPGEPSGLETLDTPPPRLIKSHLPLALLPOTLLDQKVKVYVARNPKDVA 136
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVSQADPDEIGLNMID 199
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 137 VSYHFRHMEKAHPPEPGTWSDFLEKFMAGEVSYGSHYQVQEWELSTHVPVLYLFYEDM 196
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 200 HRDLVTMVEQLARFLGVSCDKAQLBALTEHC-----HQLVDQCCNAB 242
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```

Db      :      :      :      :      :      :      :      :      :      :
197 KENPKREIQKILFEVGRSLPEETDMFQVHTSFEMKKKNPMTNTTVPQLMDHSIS---253

Qy      :      :      :      :      :      :      :      :      :      :
243 LP-VGGRGVGLWKDFTYSMNEKFPLVYKQMGKCDLTF 280

Db      :      :      :      :      :      :      :      :      :      :
254 -PPMRKGAGDKWTKTTVAQNERFPADYAERKMGCSJSF 291

```

RESULT 99

```

US-10-170-205E-17237
; Sequence 17237, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17237
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-17237

```

Query Match	30.8%;	Score 472;	DB 31;	Length 298;
Best Local Similarity	35.8%;	Pred. No. 5.3e-41;		
Matches 100;	Conservative 52;	Mismatches 101;	Indels 26;	Gaps 4;
Qy	24	GVRLPFRCGKWEIAPFVRPSDVWIVTPYKSGTSLQEVVYLVSQGADPEIGLMNID	83	
		: : : : :     : : : : :     : : : : :     : : : : :     : : : : :		
Db	20	GVPLIKYFAEALGPLSQFQARPDDLINTYKSGTWTWSQILDMIVYQGGDLKCNRAPIY	79	
		: : : : :     : : : : :     : : : : :     : : : : :     : : : : :		
Qy	84	EQLPVLEVPPOP---GLDIIKELTSPRLIKSHLPVRLPSDLHGDSKVIIVYARNPKDLV	139	
		: : : : :     : : : : :     : : : : :     : : : : :     : : : : :		
Db	80	VRVPLEVNDPGEPSGUGETLKDTPPRLIKSHLPALLPQTLLQOKKVVIVARNPKDVA	139	
		: : : : :     : : : : :     : : : : :     : : : : :     : : : : :		
Qy	140	VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFVHGVQFWEHRMDSNVLFLKYBDM	199	
		: : : : :     : : : : :     : : : : :     : : : : :     : : : : :		
Db	140	VSYVHFHRMEKAKHPGPGTWSDFLSEKFMAGESVSYGSWTQHVQEWELSKTHPVLVIFYEDM	199	
		: : : : :     : : : : :     : : : : :     : : : : :     : : : : :		
Qy	200	HRDLVTWVEQLARFLGVSCDKAQLATEHC-----HQLVDQCCNAEA	242	
		: : : : :     : : : : :     : : : : :     : : : : :     : : : : :		
Db	200	KENPKREIQKILEFVGRSLPSETWDFMVQHTSPFKEMKKNPTNYTTVPQELMDHSIS- - - -	256	
		: : : : :     : : : : :     : : : : :     : : : : :     : : : : :		
Qy	243	LP-VGRGRVGLWKDIFTVSMNKEFDLVYKQKRGKCDLTF	280	
		: : : : :     : : : : :     : : : : :     : : : : :     : : : : :		
Db	257	-PFMRKGMAGWKTTFTVAONERFDADYAEKMGAGCSLSF	294	
		: : : : :     : : : : :     : : : : :     : : : : :     : : : : :		

RESULT 100

```

RESOUR. 1000
US-09-791-537-112794
; Sequence 112794, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 112794
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-112794

```

Query Match	30.5%	Score 467;	DB 27;	Length 296;
Best Local Similarity	35.1%	Pred. No. 1.8e-40;		

	Matches	100;	Conservative	58;	Mismatches	105;	Indels	22;	Gaps	5
Qy	16	ESKYFEHGVRLP	PCFGKMEIANFVPRSDVM	VIYPKSGTSLQ	RVVYLVSG	AGAPD	75			
Db	10	QTKLKEVAGIPLQ	APTVDNNSQIQTFRAKPDLLI	CTYPKSGTTWIG	IVDMIEQNG	DVE	69			
Qy	76	EIGLWNIDEOAL	PVLEY---POP-GLDIIKE	LTSPLIKSHLP	RYFLPSDLHNGS	KVIYM	131			
Db	70	KCQRTIIQHRH	PFEIWARPPQPSGV	DKANAMPA	PRILRTHLPTQLL	PPSFWTNNCKFLVY	129			
Qy	132	ARNPKDILVSY	QYFHRSLRTMSYGT	QTQFCRRF	MNDKLG	YGSFHEHVQETFEH	RMDSNV	191		
Db	130	ARNAKDCW	SYTHFYRMSQVLP	DGCTWNEY	PETTINGK	SVGWSGFDHVKGW	WEIRDYQI	189		
Qy	192	LFLKYEDMHR	DLVTMVSQLARFL	GVSY---CDRAQL	EALTEHCHOL	VQDCCNA	EALP--	244		
Db	190	LFLFYEDVK	DPKEIKVQMFG	MKNLDEEV	VDKIVLET	SEKWK	E--NPM	TNRSTV	PKS	247
Qy	245	-----	VGRGVGLW	KDILFTVSM	MNEKFDL	IVYKQKMGK	CDLTF	280		
Db	248	VLDQSI	SPFMRKGT	VGQWKHFT	VQAQND	RDFEIVYKQKMG	GTSLN	292		

Search completed: May 17, 2006, 11:16:34  
Job time : 606 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2006, 11:17:01 ; Search time 184 Seconds  
(without alignments)  
714.962 Million cell updates/sec

Title: US-10-768-158-2

Perfect score: 1530

Sequence: 1 MAESEAETPSIPGEFESKYF.....FDLVYKQKMGKDLTFDFYL 284

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications AA Main:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1530	100.0	284	4	US-10-427-631-7
2	1530	100.0	284	4	US-10-757-262-110
3	1530	100.0	284	4	US-10-768-158-2
4	1506	98.4	284	4	US-10-205-331-46
5	480	31.4	296	4	US-10-072-012-667
6	474	31.0	1305	5	US-10-450-763-57913
7	472	30.8	295	5	US-09-854-122-21
8	472	30.8	295	5	US-10-769-507-4
9	472	30.8	295	6	US-11-033-030-21
10	467	30.5	296	4	US-10-072-012-666
11	456	29.8	295	5	US-10-206-839-3
12	456	29.8	295	5	US-10-769-507-1
13	456	29.8	295	5	US-10-769-507-3
14	456	29.8	295	5	US-10-769-507-6
15	456	29.8	295	5	US-10-287-436A-372
16	456	29.8	295	5	US-10-287-436A-1246
17	451	29.5	295	5	US-10-769-507-2
18	451	29.5	295	5	US-10-734-049A-245
19	449	29.3	302	4	US-10-072-012-586
20	449	29.3	302	5	US-10-370-715B-590
21	447	29.2	269	4	US-10-072-012-668
22	444	29.0	269	4	US-10-072-012-855
23	444	29.0	269	4	US-10-072-012-867
24	440.5	28.8	269	3	US-09-981-353-101
25	430	28.1	295	4	US-10-411-976-13
26	430	28.1	295	5	US-10-201-525-13
27	428	28.0	307	4	US-10-072-012-585
					Sequence 7, Appli
					Sequence 110, App
					Sequence 2, Appli
					Sequence 46, Appl
					Sequence 667, App
					Sequence 57913, A
					Sequence 21, Appl
					Sequence 4, Appli
					Sequence 21, Appl
					Sequence 666, App
					Sequence 3, Appli
					Sequence 1, Appli
					Sequence 3, Appli
					Sequence 6, Appli
					Sequence 372, App
					Sequence 1246, Ap
					Sequence 2, Appli
					Sequence 245, App
					Sequence 586, App
					Sequence 590, App
					Sequence 688, App
					Sequence 855, App
					Sequence 867, App
					Sequence 101, App
					Sequence 13, Appl
					Sequence 13, Appl
					Sequence 585, App

28	422	27.6	304	4	US-10-199-330-5
29	422	27.6	304	4	US-10-199-334-5
30	422	27.6	304	4	US-10-199-329-5
31	422	27.6	304	6	US-11-108-875-5
32	419.5	27.4	283	3	US-09-898-570-40
33	419.5	27.4	283	3	US-09-839-446-40
34	419.5	27.4	283	4	US-10-199-330-8
35	419.5	27.4	283	4	US-10-199-330-9
36	419.5	27.4	283	4	US-10-199-334-8
37	419.5	27.4	283	4	US-10-199-334-9
38	419.5	27.4	283	4	US-10-199-329-8
39	419.5	27.4	283	4	US-10-199-329-9
40	419.5	27.4	283	6	US-11-108-875-8
41	419.5	27.4	283	6	US-11-108-875-9
42	419.5	27.4	304	4	US-10-072-012-583
43	408.5	26.7	283	4	US-10-199-330-10
44	408.5	26.7	283	4	US-10-199-334-10
45	408.5	26.7	283	4	US-10-199-329-10
46	408.5	26.7	283	6	US-11-108-875-10
47	408.5	26.7	304	4	US-10-072-012-584
48	407	26.6	304	4	US-10-199-330-6
49	407	26.6	304	4	US-10-199-334-6
50	407	26.6	304	4	US-10-199-329-6
51	407	26.6	304	6	US-11-108-875-6
52	405	26.5	316	6	US-11-097-143-24810
53	401	26.2	304	4	US-10-199-330-4
54	401	26.2	304	4	US-10-199-334-4
55	401	26.2	304	4	US-10-199-329-4
56	401	26.2	304	4	US-10-072-012-582
57	401	26.2	304	6	US-11-108-875-4
58	399	26.1	304	4	US-10-181-108-9
59	399	26.1	304	6	US-11-079-743-9
60	397.5	26.0	295	4	US-10-072-012-206
61	396.5	25.9	305	3	US-09-898-570-26
62	396.5	25.9	305	3	US-09-839-446-26
63	385.5	25.2	297	4	US-10-381-898-4
64	383	25.0	324	4	US-10-072-012-204
65	382	25.0	76	3	US-09-989-442-149
66	381	24.9	350	5	US-10-756-149-5061
67	360.5	23.6	336	6	US-11-097-143-11343
68	359.5	23.5	346	6	US-11-097-143-14628
69	356.5	23.3	285	6	US-11-152-635-11
70	355.5	23.2	285	5	US-10-851-921-12
71	354.5	23.2	285	6	US-11-152-635-2
72	347.5	22.7	285	6	US-11-152-635-10
73	347.5	22.7	285	6	US-11-152-635-12
74	310	20.3	105	3	US-09-989-442-113
75	303.5	19.8	303	3	US-09-795-926-2
76	303.5	19.8	303	4	US-10-364-774-2
77	303.5	19.8	303	4	US-10-072-012-664
78	303.5	19.8	303	4	US-10-468-125-3
79	303.5	19.8	303	6	US-11-134-241-2
80	301	19.7	295	4	US-10-092-900A-150
81	295	19.3	313	6	US-11-097-143-11370
82	292.5	19.1	265	3	US-09-795-926-6
83	292.5	19.1	265	4	US-10-364-774-6
84	292.5	19.1	265	6	US-11-134-241-6
85	289.5	18.9	312	4	US-10-072-012-665
86	281.5	18.4	303	4	US-10-072-012-250
87	276.5	18.1	345	6	US-11-096-568A-1016
88	275.5	18.0	320	3	US-09-854-122-20
89	275.5	18.0	320	6	US-11-033-030-20
90	275	18.0	317	4	US-10-437-963-157482
91	274	17.9	282	6	US-11-096-568A-1018
92	274	17.9	286	6	US-11-096-568A-1017
93	269.5	17.6	336	4	US-10-437-963-198983
94	268.5	17.5	329	6	US-11-096-568A-7712
95	268.5	17.5	330	6	US-11-096-568A-7711
96	268.5	17.5	340	6	US-11-096-568A-7710
97	267.5	17.5	344	4	US-10-767-701-42402
98	267	17.5	348	5	US-10-739-930-7207
99	265.5	17.4	324	3	US-09-854-122-18
100	265.5	17.4	324	6	US-11-033-030-18

Sequence 5, Appli  
Sequence 5, Appli  
Sequence 5, Appli  
Sequence 40, Appli  
Sequence 40, Appli  
Sequence 8, Appli  
Sequence 9, Appli  
Sequence 8, Appli  
Sequence 9, Appli  
Sequence 8, Appli  
Sequence 9, Appli  
Sequence 583, App  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 584, App  
Sequence 6, Appli  
Sequence 6, Appli  
Sequence 6, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 582, App  
Sequence 4, Appli  
Sequence 9, Appli  
Sequence 9, Appli  
Sequence 206, App  
Sequence 26, Appl  
Sequence 4, Appli  
Sequence 204, App  
Sequence 149, App  
Sequence 5061, Ap  
Sequence 11343, A  
Sequence 14628, A  
Sequence 11, Appl  
Sequence 12, Appl  
Sequence 3, Appli  
Sequence 2, Appli  
Sequence 10, Appl  
Sequence 113, App  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 664, App  
Sequence 3, Appli  
Sequence 2, Appli  
Sequence 150, App  
Sequence 11370, A  
Sequence 6, Appli  
Sequence 6, Appli  
Sequence 6, Appli  
Sequence 665, App  
Sequence 250, App  
Sequence 250, App  
Sequence 1016, Ap  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 157482,  
Sequence 1018, Ap  
Sequence 1017, Ap  
Sequence 198983,  
Sequence 7712, Ap  
Sequence 7711, Ap  
Sequence 7710, Ap  
Sequence 42402, A  
Sequence 7207, Ap  
Sequence 18, Appl  
Sequence 18, Appl

## ALIGNMENTS

RESULT 1  
US-10-427-631-7

; Sequence 7, Application US/10427631  
; Publication No. US20030175923A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE CORPORATION; TANG, Y. Tom;  
; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;  
; APPLICANT: BAUGHN, Mariah R.; LAL, Preeti G.;  
; APPLICANT: YUE, Henry; HILLMAN, Jennifer L.;  
; APPLICANT: AZIMZAI, Yalda  
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS  
; FILE REFERENCE: PF-0592-1 DIV  
; CURRENT APPLICATION NUMBER: US/10/427,631  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: US 09/786,240  
; PRIOR FILING DATE: 1999-09-09  
; PRIOR APPLICATION NUMBER: US 60/172,220  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: US 60/155,248  
; PRIOR FILING DATE: 1998-11-04  
; PRIOR APPLICATION NUMBER: US 60/133,642  
; PRIOR FILING DATE: 1999-05-11  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PERL Program  
; SEQ ID NO 7  
; LENGTH: 284  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc feature  
; FEATURE:  
; OTHER INFORMATION: Incyte ID No. US20030175923A1 1784742CD1  
US-10-427-631-7

Query Match 100.0%; Score 1530; DB 4; Length 284;  
Best Local Similarity 100.0%; Pred. No. 1e-144;  
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAESEAETPTGPEFESKYFEFHGVRLLPPFCRGKMEEIANFPVRPSDVWIVTPKSGTSL 60  
Db 1 MAESEAETPTGPEFESKYFEFHGVRLLPPFCRGKMEEIANFPVRPSDVWIVTPKSGTSL 60  
Qy 61 LQEVVYLVSQGADPDEIGLMNIDQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Db 61 LQEVVYLVSQGADPDEIGLMNIDQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Qy 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180  
Db 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180  
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAEALTECHQLVDQCCNA 240  
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAEALTECHQLVDQCCNA 240  
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284  
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 2  
US-10-757-262-110  
; Sequence 110, Application US/10757262  
; Publication No. US20040197825A1  
; GENERAL INFORMATION:  
; APPLICANT: Karicheti, Venkateswarlu  
; APPLICANT: Silos-Santiago, Inmaculada  
; APPLICANT: Eliasof, Scott D.

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING  
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,  
; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,  
; TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43239, 44373, 51164,  
; TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,  
; TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,  
; TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,  
; TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,  
; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR  
; TITLE OF INVENTION: 55053  
; FILE REFERENCE: MPI03-007P1RNMNMIM  
; CURRENT APPLICATION NUMBER: US/10/757,262  
; CURRENT FILING DATE: 2004-01-14  
; PRIOR APPLICATION NUMBER: US 60/440,318  
; PRIOR FILING DATE: 2003-01-15  
; PRIOR APPLICATION NUMBER: US 60/444,783  
; PRIOR FILING DATE: 2003-02-04  
; PRIOR APPLICATION NUMBER: US 60/457,901  
; PRIOR FILING DATE: 2003-03-27  
; PRIOR APPLICATION NUMBER: US 60/468,775  
; PRIOR FILING DATE: 2003-05-08  
; PRIOR APPLICATION NUMBER: US 60/471,614  
; PRIOR FILING DATE: 2003-05-19  
; PRIOR APPLICATION NUMBER: US 60/478,742  
; PRIOR FILING DATE: 2003-06-16  
; PRIOR APPLICATION NUMBER: US 60/488,529  
; PRIOR FILING DATE: 2003-07-18  
; PRIOR APPLICATION NUMBER: US 60/491,156  
; PRIOR FILING DATE: 2003-07-30  
; PRIOR APPLICATION NUMBER: US 60/499,594  
; PRIOR FILING DATE: 2003-09-02  
; PRIOR APPLICATION NUMBER: US 60/506,332  
; PRIOR FILING DATE: 2003-09-26  
; NUMBER OF SEQ ID NOS: 136  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 110  
; LENGTH: 284  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-757-262-110  
Query Match 100.0%; Score 1530; DB 4; Length 284;  
Best Local Similarity 100.0%; Pred. No. 1e-144;  
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAESEAETPTGPEFESKYFEFHGVRLLPPFCRGKMEEIANFPVRPSDVWIVTPKSGTSL 60  
Db 1 MAESEAETPTGPEFESKYFEFHGVRLLPPFCRGKMEEIANFPVRPSDVWIVTPKSGTSL 60  
Qy 61 LQEVVYLVSQGADPDEIGLMNIDQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Db 61 LQEVVYLVSQGADPDEIGLMNIDQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120  
Qy 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180  
Db 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180  
Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAEALTECHQLVDQCCNA 240  
Db 181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAEALTECHQLVDQCCNA 240  
Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284  
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284  
RESULT 3  
US-10-768-158-2  
; Sequence 2, Application US/10768158  
; Publication No. US20040204359A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Silos-Santiago, Inmaculada

```
; APPLICANT: Karicheti, Venkateswarlu
; APPLICANT: Eliasof, Scott D.
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS FOR TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 16386, 15402, 21165, 1423,
; TITLE OF INVENTION: 636, 12303, 21425, 27410, 38554, 38555, 55063, 57145, 59914,
; TITLE OF INVENTION: 94921, 16852, 33260, 58573, 30911, 85913, 14303, 16816,
; TITLE OF INVENTION: 17827 OR 32620
; FILE REFERENCE: MPI03-012P1RNMIM
; CURRENT APPLICATION NUMBER: US/10/768,158
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: US 60/444,781
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/452,291
; PRIOR FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: US 60/454,540
; PRIOR FILING DATE: 2003-03-13
; PRIOR APPLICATION NUMBER: US 60/478,805
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/491,048
; PRIOR FILING DATE: 2003-07-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-10-768-158-2

Query Match      100.0%; Score 1530; DB 4; Length 284;
Best Local Similarity 100.0%; Pred. No.1e-144;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAESEATPTPGEPESKYFFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPTPGEPESKYFFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60

Qy 61 LQEVVYLSQAGADPEIGLMNIDEQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLSQAGADPEIGLMNIDEQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYMARNPKDLVSVYQPHRSRLTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYMARNPKDLVSVYQPHRSRLTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180

Qy 181 EFWEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCKAQLALETCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCKAQLALETCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 4
US-10-205-331-46
; Sequence 46, Application US/10205331
; Publication No. US20040058326A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Allstair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnoch, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018199
; CURRENT APPLICATION NUMBER: US/10/205,331
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 284
```

```
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Sulfotransferase-like protein
US-10-205-331-46

Query Match      98.4%; Score 1506; DB 4; Length 284;
Best Local Similarity 97.9%; Pred. No.2.6e-142;
Matches 278; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAESEATPTPGEPESKYFFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEATPTPGEPESKYFFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPKSGTSL 60

Qy 61 LQEVVYLSQAGADPEIGLMNIDEQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LQEVVYLSQAGADPEIGLMNIDEQLPVLEYPOQGLDIIKELTSPRLIKSHLPYRFLPSD 120

Qy 121 LHNGDSKVIYMARNPKDLVSVYQPHRSRLTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Db 121 LHNGDSKVIYMARNPKDLVSVYQPHRSRLTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180

Qy 181 EFWEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCKAQLALETCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCKAQLALETCHQLVDQCCNA 240

Qy 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 5
US-10-072-012-667
; Sequence 667, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
```

```
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 667
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-072-012-667

Query Match      31.4%; Score 480; DB 4; Length 296;
Best Local Similarity 35.6%; Pred. No. 2.7e-39;
Matches 103; Conservative 57; Mismatches 107; Indels 22; Gaps 5;

Qy 16 ESKYFEFHGVLPPPCRCGKMBEIANFPVRPSDVMTVTYPKSGTLLQEVVYLVSGGADPD 75
Db 10 QTKUKEVAGIFLQAPTVDNWNQIQITFEAKPDLLICTTPKSGTTWQIVDMIQNGDVE 69

Qy 76 EIGLMNIDEQLVLEY---POP-GLDIIKELTSPRLIKSHLPYRPLPSDLHNGDSKVIYM 131
Db 70 KCRRTIIQHRHPFIWARPPQPSGVDKANEMPAPRILRTHLPTQLLPSPFWTNCKFLYV 129

Qy 132 ARNPKDLVVSYYQFHRSLRTWSYRGTFQFCRFRFNDKLGYSWFHVEHVEHMRMDSNV 191
Db 130 ARNAKDCMVSYHYFHRMSQVLPEPQTWDEYFETFTINGKVSFMDHVKGWMEIRDKYQI 189

Qy 192 LFLKYEDMHRDLVTMVEQLARFLGVS-----CDKAQLEALTEHCHOLVDQCCNAEALP-- 244
Db 190 LFLFYEDMKRNPKEHIQKVMQFMGKNLDEVDVVKLVLETSFEKMKKE--NPNMNRSTAPKS 247

Qy 245 -----VGRGRVGLMKDIFTVSMNEKFDLVYKQKMGKCDLTFDFYL 284
Db 248 ILDQISPFMRKGTGDMKNHFTVAQNERFDEIYKQKMGRTSLNFSMEL 296

RESULT 6
US-10-450-763-57913
; Sequence 57913, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 57913
; LENGTH: 1305
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: {649}...(695)
; OTHER INFORMATION: Serine/threonine specific protein phosphatases
; OTHER INFORMATION: domain identified by eMATRIX, accession number BL00125B, p-value=
; OTHER INFORMATION: 1.000e-40, raw score of 21.48
; FEATURE:
```

```
; NAME/KEY: DOMAIN
; LOCATION: {593}...(850)
; OTHER INFORMATION: Ser/Thr protein phosphatase domain identified by Pfam,
; OTHER INFORMATION: accession name Stposphatase, E-value=2.6e-196, Pfam score of 665
US-10-450-763-57913

Query Match      31.0%; Score 474; DB 5; Length 1305;
Best Local Similarity 33.5%; Pred. No. 8.6e-38;
Matches 110; Conservative 54; Mismatches 110; Indels 54; Gaps 8;

Qy 3 ESEAEETSTP-----GEPESKYFEFHGVL-----PP--FCRG-----K 34
Db 265 EEGREQPVPTVTTPIILSPCRQGENRSGRSLRNNMELIQDTSRPPLEVYKGVPLIKYPAEA 324

Qy 35 MEEIANFPVRPSDVMTVTYPKSGTLLQEVVYLVSGGADPDDEIGLMNIDEQLPVLEYPOP 94
Db 325 LGPLQSFQARPDLLINTYPKSGTTWVSQILDMTYQGGDLKCNRAPIYVRVPFLEVNDP 384

Qy 95 ----GLDIIKELTSPRLIKSHLPYRPLPSDLHNGDSKVIYMARNPKDLVVSYYQFHRSLR 150
Db 385 GEPGSLTLDKDTPPRLIKSHLPALLPQLDQKVVYVARNPKDVAVSYHFRHMEK 444

Qy 151 TMSYRGTFQFCRFRFNDKLGYSWFHVEHMRMDSNVLFKYEDMHRDLVTMVEQL 210
Db 445 AHPEPGTWSFLEKFMAGEVSYSYQVQEWELSRTHPVLYLFYEDMKENPKREIQKI 504

Qy 211 ARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAEALP-VGRGRVGL 252
Db 505 LEFVGRSLPEETPMVQHTSFKETKKNPMNTYTPQELMDHSIS----PFMRKGMAGD 560

Qy 253 WKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 561 WKTFTVAQNERFDADVAEKMGACSLSF 588

RESULT 7
US-09-854-122-21
; Sequence 21, Application US/09854122
; Patent No. US20020016980A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTE, RANDALL S.
; APPLICANT: SMITH, ROBERT
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
; FILE REFERENCE: PHA-007.01
; CURRENT APPLICATION NUMBER: US/09/854,122
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/202,529
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-854-122-21

Query Match      30.8%; Score 472; DB 3; Length 295;
Best Local Similarity 35.8%; Pred. No. 1.7e-38;
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;

Qy 24 GVRLPFFCRGKMBEIANFPVRPSDVMTVTYPKSGTLLQEVVYLVSGGADPDDEIGLMNID 83
Db 17 GVPLIKYFAEALGFQSFQARPDLLINTYPKSGTTWVSQILDMTYQGGDLKCNRAPIY 76

Qy 84 EQLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRPLPSDLHNGDSKVIYMARNPKDLV 139
Db 77 VRVPFLEVNDPGEPSGLETLKDTPPRLIKSHLPALLPQLDQKVVYVARNPKDV 136

Qy 140 VSYQFHRSLRTMYRGTFQFCRFRFNDKLGYSWFHVEHMRMDSNVLFKYEDM 199
Db 137 VSYHFRHMEKAPPEPGTWSFLEKFMAGEVSYSYQVQEWELSRTHPVLYLFYEDM 196

Qy 200 HRDLVTMVEQLARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAE 242
```



Db 197 KENPKREIQKILEFVGRSLPEETDMFVQHTSFKEMKKNPMTNTYTPQELMDHSIS--- 253  
Qy 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
Db 254 -PFMRKGMAGDWKTTFTTVAQNERFDADYAEKMGAGCSLSF 291

## RESULT 8

US-10-769-507-4  
; Sequence 4, Application US/10769507  
; Publication No. US20050019788A1  
; GENERAL INFORMATION:  
; APPLICANT: Squires, James E.  
; TITLE OF INVENTION: Genetic Markers for Skatole Metabolism  
; FILE REFERENCE: P06153US03  
; CURRENT APPLICATION NUMBER: US/10/769,507  
; CURRENT FILING DATE: 2004-01-30  
; PRIOR APPLICATION NUMBER: 10/024,628  
; PRIOR FILING DATE: 2001-11-23  
; PRIOR APPLICATION NUMBER: 09/288,037  
; PRIOR FILING DATE: 1999-04-08  
; PRIOR APPLICATION NUMBER: 60/081,037  
; PRIOR FILING DATE: 1998-04-08  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 4  
; LENGTH: 295  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-769-507-4

Query Match 30.8%; Score 472; DB 5; Length 295;  
Best Local Similarity 35.8%; Pred. No. 1.7e-38;  
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;  
Qy 24 GVRLLPPFCRGKMEIANPVRPSDWIVTYPKSGTSLLOEVVYLVSQADPDEIGLMMID 83  
Db 17 GVPLIKYPAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDMYQGGDLEKCNRAPIY 76  
Qy 84 EOLPVLVEYPOP---GLDIIKELTSPRLIKSHLPVRLPSDLHNGDSKVIYMARNP KD LV 139  
Db 77 VRVFLVNDPGEPSGLETLDTPPRLIKSHLPALLPQTLLOKVKVYVARNPKDVA 136  
Qy 140 VSYTQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVFWEHRMDSNVLFKLYEDM 199  
Db 137 VSYVHFHRMEKAHPPEPGTWSDFLEKFMAGEVSYSGWYQHVQEWELSRTHPVLVLYFYEDM 196  
Qy 200 HRDLVTWVEQLARFLGVSCDKAQLAETHC-----HQLVDQCCNAEA 242  
Db 197 KENPKREIQKILEFVGRSLPEETDMFVQHTSFKEMKKNPMTNTYTPQELMDHSIS--- 253  
Qy 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
Db 254 -PFMRKGMAGDWKTTFTTVAQNERFDADYAEKMGAGCSLSF 291

## RESULT 9

US-11-033-030-21  
; Sequence 21, Application US/11033030  
; Publication No. US20060053510A1  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTE, RANDALL S.  
; APPLICANT: SMITH, ROBERT  
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA  
; FILE REFERENCE: PHA-007.01  
; CURRENT APPLICATION NUMBER: US/11/033,030  
; CURRENT FILING DATE: 2005-01-11  
; PRIOR APPLICATION NUMBER: US/09/854,122  
; PRIOR FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: 60/202,529  
; PRIOR FILING DATE: 2000-05-10

; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: Patent in ver. 2.1  
; SEQ ID NO 21  
; LENGTH: 295  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-033-030-21

Query Match 30.8%; Score 472; DB 6; Length 295;  
Best Local Similarity 35.8%; Pred. No. 1.7e-38;  
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;  
Qy 24 GVRLLPPFCRGKMEIANPVRPSDWIVTYPKSGTSLLOEVVYLVSQADPDEIGLMMID 83  
Db 17 GVPLIKYPAEALGPLQSFQARPDLLINTYPKSGTTWVSQILDMYQGGDLEKCNRAPIY 76  
Qy 84 EOLPVLVEYPOP---GLDIIKELTSPRLIKSHLPVRLPSDLHNGDSKVIYMARNP KD LV 139  
Db 77 VRVFLVNDPGEPSGLETLDTPPRLIKSHLPALLPQTLLOKVKVYVARNPKDVA 136  
Qy 140 VSYTQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVFWEHRMDSNVLFKLYEDM 199  
Db 137 VSYVHFHRMEKAHPPEPGTWSDFLEKFMAGEVSYSGWYQHVQEWELSRTHPVLVLYFYEDM 196  
Qy 200 HRDLVTWVEQLARFLGVSCDKAQLAETHC-----HQLVDQCCNAEA 242  
Db 197 KENPKREIQKILEFVGRSLPEETDMFVQHTSFKEMKKNPMTNTYTPQELMDHSIS--- 253  
Qy 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
Db 254 -PFMRKGMAGDWKTTFTTVAQNERFDADYAEKMGAGCSLSF 291

## RESULT 10

US-10-072-012-666  
; Sequence 666, Application US/10072012  
; Publication No. US20040033493A1  
; GENERAL INFORMATION:  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Li, Li  
; APPLICANT: Gangolli, Esha  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Anderson, David W.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Taupier Jr, Raymond J.  
; APPLICANT: Gusev, Vladimir Y.  
; APPLICANT: Colman, Steven D.  
; APPLICANT: Wolenc, Adam R.  
; APPLICANT: Pena, Carol E. A  
; APPLICANT: Furtak, Katarzyna  
; APPLICANT: Grosse, William M.  
; APPLICANT: Alsobrook II, John P.  
; APPLICANT: Lepley, Denise M.  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Burgess, Catherine E.  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-258  
; CURRENT APPLICATION NUMBER: US/10/072,012  
; CURRENT FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: 60/265,102  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 60/265,514  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,517  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,412  
; PRIOR FILING DATE: 2001-01-31

```
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 666
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-072-012-666

Query Match      30.5%; Score 467; DB 4; Length 296;
Best Local Similarity 35.1%; Pred. No. 5.5e-38;
Matches 100; Conservative 58; Mismatches 105; Indels 22; Gaps 5;

Qy 16 ESKYFEHGVLPFCRCGMEEIANFPVRPSDVMIVTPKSGTSLLOEVVYLVSGQADPD 75
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 76 EIGLWNIDEQPLVLEY---POP-GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYM 131
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 70 KCQRTIQRHPFTEWARPPGSGVDKANAPAFILTHLPTQLLPSPFWTNCKELYV 129
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 132 ARNPKDLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVEFWEHRMDSNV 191
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 130 ARNAKDCMVSYHFYRMSQVLPDGTWNEYPTETFGKVSNGSFDHVKGWWEIRDYQI 189
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 192 LFLKYEDMHRDLVTMVEQLARFLGVS-----CDKAQLEALTEHCQLVDQCCNAELP-- 244
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 190 LFLFYEDVKRDPKREIQVMQFMGNLDEEVVDKIVLETSEFKMKE--NPMTNRTSVPKS 247
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 245 -----VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 248 VLDQISFPFMKGTGVDKXNHFTVAQNDRFDEIYKQMGGTSLNP 292
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
US-10-206-839-3
; Sequence 3, Application US/10206839
; Publication No. US2003009977A1
; GENERAL INFORMATION:
; APPLICANT: Guida, Marco
; APPLICANT: Kurth, Janice
; TITLE OF INVENTION: Genotyping Human Phenol Sulfotransferase
; FILE OF INVENTION: (STP2)
; TITLE REFERENCE: 4389-6 (formerly SEQ-16P)
; CURRENT APPLICATION NUMBER: US/10/206,839
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 09/328,174
; PRIOR FILING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 295
; TYPE: PRT
; ORGANISM: H. sapiens
US-10-206-839-3

Query Match      29.8%; Score 456; DB 4; Length 295;
Best Local Similarity 33.7%; Pred. No. 7e-37;
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;

Qy 24 GVRLPPFCRCGMEEIANFPVRPSDVMIVTPKSGTSLLOEVVYLVSGQADPDDEIGLWNID 83
```

```
Db 17 GVPLIKYFAEALGLOFSQFARPDLLISTYPSKGTTWVSQILDMIYOGGLEKCHRAPIF 76
Qy 84 EQLPVLVEYPOP---GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNPXDLV 139
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 77 MRVPFLFKVPFGIPSGMETLKNTPAPRLKTLPLALLPQTLDDQKVKVYVARNAXDVA 136
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 140 VSYIQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVEFWEHRMDSNVLEFLKYEDM 199
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 137 VSYHYFYHMAKVYPHPGCTWESFLEKFMAGEVSYGSWYQHVQWMELSRTHFVLYLFYEDM 196
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 200 HRDLVTMVEQLARFLGVSCKAOLEALTEHC-----HQLVDQCCNAEA 242
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 197 KENPKREIQKILEFVGRSLPEETVDLMVENTSPKEMKKNPMNTYTVRRFMDHSIS--- 253
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 254 -PFMRKGMAGDWKTTFTVAQNERFDADYAEKMGAGCSLSF 291
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
US-10-769-507-1
; Sequence 1, Application US/10769507
; Publication No. US20050019788A1
; GENERAL INFORMATION:
; APPLICANT: Squires, James E.
; APPLICANT: Lin, Zhihong E.
; TITLE OF INVENTION: Genetic Markers for Skatole Metabolism
; FILE REFERENCE: P06153US03
; CURRENT APPLICATION NUMBER: US/10/769,507
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: 10/024,628
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: 09/288,037
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/081,037
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 295
; TYPE: PRT
; ORGANISM: porcine
US-10-769-507-1

Query Match      29.8%; Score 456; DB 5; Length 295;
Best Local Similarity 35.6%; Pred. No. 7e-37;
Matches 98; Conservative 51; Mismatches 108; Indels 18; Gaps 3;

Qy 24 GVRLPPFCRCGMEEIANFPVRPSDVMIVTPKSGTSLLOEVVYLVSGADPDDEIGLWNID 83
Db 17 GVPLIKYFAEALGLOFSQFARPDLLISTYPSKGTTWVSSEILDIIYOGGDLQKQCRAPIF 76
Qy 84 EQLPVLVEYPOP---GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNPXDLV 139
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 77 VRVPFLFKVPFCPTGFELKDTAPAPRLKTLPLTLLPQTLDDQKVKVYVARNAXDVA 136
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 140 VSYIQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVEFWEHRMDSNVLEFLKYEDM 199
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 137 VSYHYFYHMAKVYPNPCTWDSFLEDFMAGEVSYGSWYQHVQWMELSRTHFVLYLFYEDM 196
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 200 HRDLVTMVEQLARFLGVSCKAOLEALTEHC-----HQLVDQCCNAELP-----V 245
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 197 KENPKREIQKILEFVGRSLPEETVEDIVQHTSFOEMKKNAMNTYRTLPDILLDHSISAFM 256
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 246 GRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 257 RKGITGDKSTFTTVAQNERFEADYAEKMGAGNLR 291
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
US-10-769-507-3
; Sequence 3, Application US/10769507
```

```

; Publication No. US20050019788A1
; GENERAL INFORMATION:
; APPLICANT: Squires, James E.
; APPLICANT: Lin, Zhihong E.
; TITLE OF INVENTION: Genetic Markers for Skatole Metabolism
; FILE REFERENCE: P06153US03
; CURRENT APPLICATION NUMBER: US/10/769,507
; PRIOR FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: 10/024,628
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: 09/288,037
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/081,037
; PRIOR FILING DATE: 1998-04-08
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-769-507-3

Query Match      29.8%; Score 456; DB 5; Length 295;
Best Local Similarity 33.7%; Pred. No. 7e-37;
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLLOEVVYLVVSQGDADPEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSFQARPDLLISTYPKSGTTWVSQILDMIYQGGDLKCHRAPIF 76
QY 84 EQLPVLEYPOP-----GLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYIYARNPKDLV 139
DB 77 MRVPFLEFKIPGPGSMETLKNTPAPRLKTLHLPLTLLPQTLDDQKVVVYVARNAKOVA 136
QY 140 VSYIQFHRSRLTMSYRGTFQFCRRFMDKLGYSWFEHVOEFWEHRMDSNVFLFLKYEDM 199
DB 137 VSYIHFYHMAKVYPHGGTWESFLEKFMAGEVSYGSWYQHVQEWELSRTHPVLYLFYEDM 196
QY 200 HRDLVTMVEQLARFLGVSQDKAQLAETHC-----HQLVDQCCNAEA 242
DB 197 KENPKREIQKILEFVGRSLPEETVDMVEHTSFKEMKNPMTNYTTVTRREFMDHSIS 253
QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 254 -PFMRKGMAGDWKTTFTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 14
US-10-769-507-6
; Sequence 6, Application US/10769507
; Publication No. US20050019788A1
; GENERAL INFORMATION:
; APPLICANT: Squires, James E.
; APPLICANT: Lin, Zhihong E.
; TITLE OF INVENTION: Genetic Markers for Skatole Metabolism
; FILE REFERENCE: P06153US03
; CURRENT APPLICATION NUMBER: US/10/769,507
; PRIOR FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: 10/024,628
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: 09/288,037
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/081,037
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 295
; TYPE: PRT
; ORGANISM: porcine
US-10-769-507-6

Query Match      29.8%; Score 456; DB 5; Length 295;
Best Local Similarity 33.7%; Pred. No. 7e-37;
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLLOEVVYLVVSQGDADPEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSFQARPDLLISTYPKSGTTWVSQILDMIYQGGDLKCHRAPIF 76
QY 84 EQLPVLEYPOP-----GLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYIYARNPKDLV 139
DB 77 MRVPFLEFKIPGPGSMETLKNTPAPRLKTLHLPLTLLPQTLDDQKVVVYVARNAKOVA 136
QY 140 VSYIQFHRSRLTMSYRGTFQFCRRFMDKLGYSWFEHVOEFWEHRMDSNVFLFLKYEDM 199
DB 137 VSYIHFYHMAKVYPHGGTWESFLEKFMAGEVSYGSWYQHVQEWELSRTHPVLYLFYEDM 196
QY 200 HRDLVTMVEQLARFLGVSQDKAQLAETHC-----HQLVDQCCNAEA 242
DB 197 KENPKREIQKILEFVGRSLPEETVDMVEHTSFKEMKNPMTNYTTVTRREFMDHSIS 253
QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 254 -PFMRKGMAGDWKTTFTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 15
US-10-287-436A-372
; Sequence 372, Application US/10287436A
; Publication No. US20050202421A1
; GENERAL INFORMATION:
; APPLICANT: CHILDREN'S HOSPITAL MEDICAL CENTER
; TITLE OF INVENTION: METHOD FOR DIAGNOSIS AND TREATMENT OF
; FILE REFERENCE: 10872.514696
; CURRENT APPLICATION NUMBER: US/10/287,436A
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/336,220
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 1446
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 372
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-436A-372

Query Match      29.8%; Score 456; DB 5; Length 295;
Best Local Similarity 33.7%; Pred. No. 7e-37;
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLLOEVVYLVVSQGDADPEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSFQARPDLLISTYPKSGTTWVSQILDMIYQGGDLKCHRAPIF 76
QY 84 EQLPVLEYPOP-----GLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYIYARNPKDLV 139
DB 77 MRVPFLEFKIPGPGSMETLKNTPAPRLKTLHLPLTLLPQTLDDQKVVVYVARNAKOVA 136
QY 140 VSYIQFHRSRLTMSYRGTFQFCRRFMDKLGYSWFEHVOEFWEHRMDSNVFLFLKYEDM 199
DB 137 VSYIHFYHMAKVYPHGGTWESFLEKFMAGEVSYGSWYQHVQEWELSRTHPVLYLFYEDM 196
QY 200 HRDLVTMVEQLARFLGVSQDKAQLAETHC-----HQLVDQCCNAEA 242
DB 197 KENPKREIQKILEFVGRSLPEETVDMVEHTSFKEMKNPMTNYTTVTRREFMDHSIS 253
QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 254 -PFMRKGMAGDWKTTFTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 16
US-10-287-436A-1246
; Sequence 1246, Application US/10287436A
```

```
; Publication No. US20050202421A1
; GENERAL INFORMATION:
; APPLICANT: CHILDREN'S HOSPITAL MEDICAL CENTER
; TITLE OF INVENTION: METHOD FOR DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS
; FILE REFERENCE: 10872.514696
; CURRENT APPLICATION NUMBER: US/10/287,436A
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/336,220
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 1446
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1246
; LENGTH: 295
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-287-436A-1246

Query Match      29.8%; Score 456; DB 5; Length 295;
Best Local Similarity 33.7%; Pred. No. 7e-37;
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLLPFCRCGKMEETANFVRPSDVWIVTPKSGTSLLEQVYVLSQGADPDEIGLMNID 83
Db 17 GVPLIKYFAEALGPLQSFQARPDLLISTYPKSGTTWVSQILDMYQGGDLEKCHRAPIF 76
;
QY 84 BOLPVLEYPOP-GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNP KD LV 139
;
Db 77 MRVPFLEFKAPGPSGMEETLKNTAPRLLKTHLPLALLPQTLDDQKVKVYVARNAKDVA 136
;
QY 140 VSYQYFHRSLRMTSYRGTFQFCRRFMNDKLGYSWEHVEHVEHFMDSNVFLFKYEDM 199
;
Db 137 VSYHYFHYHMAKVHPPEPGTWDSEFLKFMVGEVSGSYQHVQEWELSRTHPVLYLFYEDM 196
;
QY 200 HRDLVTMVEQLARFLGVSCDKAQLAEALTEHC-----HQLVDQCCNAEALP- 245
;
Db 197 KENPKREIQKILEFVGHSLPEETVDFMVQHTSFKEKMKKNPMTNTTVPQEFMDHSISPFM 256
;
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 257 RKG MAGDWKTTTVAQNERFDADYAEK MAGCSLSF 291
;
RESULT 18
US-10-734-049A-245
; Sequence 245, Application US/10734049A
; Publication No. US20050042624A1
; GENERAL INFORMATION:
; APPLICANT: ITOH, Kyogo
; TITLE OF INVENTION: TUMOR ANTIGEN
; FILE REFERENCE: Q-78382
; CURRENT APPLICATION NUMBER: US/10/734,049A
; CURRENT FILING DATE: 2003-12-12
; PRIOR APPLICATION NUMBER: PCT/JP02/05799
; PRIOR FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: JP 2001/177058
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: JP 2001/250728
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 408
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 245
; LENGTH: 295
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-734-049A-245

Query Match      29.5%; Score 451; DB 5; Length 295;
Best Local Similarity 33.8%; Pred. No. 2.2e-36;
Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;

QY 24 GVRLLPFCRCGKMEETANFVRPSDVWIVTPKSGTSLLEQVYVLSQGADPDEIGLMNID 83
Db 17 GVPLIKYFAEALGPLQSFQARPDLLISTYPKSGTTWVSQILDMYQGGDLEKCHRAPIF 76
;
QY 84 BOLPVLEYPOP-GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNP KD LV 139
;
Db 77 MRVPFLEFKAPGPSGMEETLKNTAPRLLKTHLPLALLPQTLDDQKVKVYVARNAKDVA 136
;
QY 140 VSYQYFHRSLRMTSYRGTFQFCRRFMNDKLGYSWEHVEHVEHFMDSNVFLFKYEDM 199
;
Db 137 VSYHYFHYHMAKVHPPEPGTWDSEFLKFMVGEVSGSYQHVQEWELSRTHPVLYLFYEDM 196
;
QY 200 HRDLVTMVEQLARFLGVSCDKAQLAEALTEHC-----HQLVDQCCNAEALP- 245
;
Db 197 KENPKREIQKILEFVGHSLPEETVDFMVQHTSFKEKMKKNPMTNTTVPQEFMDHSISPFM 256
;
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 257 RKG MAGDWKTTTVAQNERFDADYAEK MAGCSLSF 291
;
RESULT 19
US-10-072-012-586
; Sequence 586, Application US/10072012
```

```

; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkete, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grobse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 586
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-072-012-586

Query Match      29.3%; Score 449; DB 4; Length 302;
Best Local Similarity 36.2%; Pred. No. 3.6e-36;
Matches 96; Conservative 52; Mismatches 99; Indels 18; Gaps 3;

Qy 36 BEIANFVRPSDWMVTVPKSGTSLQEVVYLVSGADPDEIGLMNIDBQLPVLEYPOP- 94
Db 36 DKWNFOAKPDLLISTYPKAGTTWTQEIIVELIQNEGVDEKSKRAPHQRPFFLEMKIPS 95
Qy 95 ---GLDIIKELTSPRLIKSHLPYRPLPSDLHNGDSKVLYMARNPKDLVSYVYQFHSRLT 151
Db 96 LGSGLQEAHAMPSPRIILKTHLPFLLPSPLEKNCKIIVARNPKDNMVSYYHFORMNKA 155
Qy 152 MSYRGTFQFCRRFMNDKLGYSWFHVEHWMDSNVLFKYEDMHRDLVTMVEOLA 211
Db 156 LPAPGTWEEYFETFLAGVKGWSWHEHVKGWWEAKDKHRLILFYEDMKKPKHIEIQKLA 215
Qy 212 RFLGVSCDCAQLEALTEHCHQLV---DQCCNAEALP-----VGRGRVGLWKDIF 257
Db 216 EFIGKLLDDKVLDKIVHYTSFDMVKQNPANYSSIPAEIMDHSISPPFMRKGAVGDKKHF 275
Qy 258 TVSMNEKFDLVYKQMGKCDLTFDF 282
Db 276 TVAQNERFDEYDKKMTDTRLTFHF 300

RESULT 21
US-10-072-012-668
; Sequence 668, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkete, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha

Qy 212 RFLGVSCDCAQLEALTEHCHQLV---DQCCNAEALP-----VGRGRVGLWKDIF 257
Db 216 EFIGKLLDDKVLDKIVHYTSFDMVKQNPANYSSIPAEIMDHSISPPFMRKGAVGDKKHF 275
Qy 258 TVSMNEKFDLVYKQMGKCDLTFDF 282
Db 276 TVAQNERFDEYDKKMTDTRLTFHF 300

US-10-370-715B-590
; Sequence 590, Application US/10370715B
; Publication No. US20040258678A1
; GENERAL INFORMATION:
; Patin Docket Preview
; APPLICANT: BODARY SARAH C.
; APPLICANT: CLARK, HILLARY
; APPLICANT: BRISDELL, HUNTE
; APPLICANT: JACKMAN, JANET
; APPLICANT: SCHOENFELD, JILL R.
; APPLICANT: WILLIAMS, P. MICKEY
; APPLICANT: WOOD, WILLIAM I.
; APPLICANT: WU, THOMAS D.
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune
; TITLE OF INVENTION: Related Diseases
; FILE REFERENCE: P1948R1-US
; CURRENT APPLICATION NUMBER: US/10/370,715B
; CURRENT FILING DATE: 2003-02-21
; NUMBER OF SEQ ID NOS: 742
; SEQ ID NO 590
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapien
; US-10-370-715B-590

Query Match      29.3%; Score 449; DB 5; Length 302;
Best Local Similarity 36.2%; Pred. No. 3.6e-36;
Matches 96; Conservative 52; Mismatches 99; Indels 18; Gaps 3;

Qy 36 BEIANFVRPSDWMVTVPKSGTSLQEVVYLVSGADPDEIGLMNIDBQLPVLEYPOP- 94
Db 36 DKWNFOAKPDLLISTYPKAGTTWTQEIIVELIQNEGVDEKSKRAPHQRPFFLEMKIPS 95
Qy 95 ---GLDIIKELTSPRLIKSHLPYRPLPSDLHNGDSKVLYMARNPKDLVSYVYQFHSRLT 151
Db 96 LGSGLQEAHAMPSPRIILKTHLPFLLPSPLEKNCKIIVARNPKDNMVSYYHFORMNKA 155
Qy 152 MSYRGTFQFCRRFMNDKLGYSWFHVEHWMDSNVLFKYEDMHRDLVTMVEOLA 211
Db 156 LPAPGTWEEYFETFLAGVKGWSWHEHVKGWWEAKDKHRLILFYEDMKKPKHIEIQKLA 215
Qy 212 RFLGVSCDCAQLEALTEHCHQLV---DQCCNAEALP-----VGRGRVGLWKDIF 257
Db 216 EFIGKLLDDKVLDKIVHYTSFDMVKQNPANYSSIPAEIMDHSISPPFMRKGAVGDKKHF 275
Qy 258 TVSMNEKFDLVYKQMGKCDLTFDF 282
Db 276 TVAQNERFDEYDKKMTDTRLTFHF 300

RESULT 21
US-10-072-012-668
; Sequence 668, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkete, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
```





```
Db 196 DMKNPKKEIKKIRFLKLEKNLNDLRIIHTTSFVVMKDNPLVNYTHLPTTVMDSKSP 255
Qy 245 -VGRGRVGLWKDIFTVSNNEKFDLVYKQMGKCDLTF 280
      : : | | | | | | | | | | | | | | | | | | | |
Db 256 PMRGKTAGDKNKFYVAQNEKFDALYETEMSKTALQF 292

RESULT 25
US-10-411-976-13
; Sequence 13, Application US/10411976
; Publication No. US20030170849A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Kevin L.
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES
; FILE REFERENCE: 5864,027
; CURRENT APPLICATION NUMBER: US/10/411,976
; PRIOR FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 09/785,343
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: estrogen sulfotransferase
US-10-411-976-13

Query Match 28.1%; Score 430; DB 4; Length 295;
Best Local Similarity 33.0%; Pred. No. 2.8e-34;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

Qy 14 EFESKYFEFHGVRLPPFCRGKMEBIANFPVRPSDVMTVTPKSGTSLLEVVYLVSQAD 73
      : : | | | | | | | | | | | | | | | | | | | |
Db 7 EYEVFGFGRGLMDKRFYKYWEDVENFLARPDDLVIATPKSGTTWISVVYMIYKEGD 66

Qy 74 PDEIGLMNIDQLPVLEYPOP-----GLDI IKELTSPRLIKSHLPYRFLPSDLHNGDSKVI 129
      : : | | | | | | | | | | | | | | | | | | | |
Db 67 VEKCKEDAIFNRIPYLECRNEDLINGIKQLKEKESPRIVKTHLPKVLPA SFWEKNCCKMI 126

Qy 130 YMARNPDLVSYVYQFHRSLRTMSYRGTFQFCFRFMNDKLGYSWFEHVSQFWEHRMDS 189
      : : | | | | | | | | | | | | | | | | | | | |
Db 127 YLCRNADKDVAVSYVYFLMLMITSYPNPKSFSEFVEKFMQGVPGSWYDHWKAWWEKSKNS 186

Qy 190 NVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCH-----QLVDQ 236
      : : | | | | | | | | | | | | | | | | | | | |
Db 187 RVLFMFYEDMKEDIRREVVKLIEFLERKPSAELVDRIIQTSTFOEMKNPSTNYTMPEE 246

Qy 237 CCNAEALP-VGRGRVGLWKDIFTVSNNEKFDLVYKQMGKCDLTF 280
      : : | | | | | | | | | | | | | | | | | | | |
Db 247 MNQKVSFPMRKGIIGDNKNHFPPEALRERFDEHYKQMKDCTVKF 291

RESULT 26
US-10-201-525-13
; Sequence 13, Application US/10201525
; Publication No. US20060009631A1
; GENERAL INFORMATION:
; APPLICANT: Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5827,005
; CURRENT APPLICATION NUMBER: US/10/201,525
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/785,343
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: PCT/US99/16750
; PRIOR FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-201-525-13

Query Match 28.1%; Score 430; DB 4; Length 295;
Best Local Similarity 33.0%; Pred. No. 2.8e-34;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

Qy 14 EFESKYFEFHGVRLPPFCRGKMEBIANFPVRPSDVMTVTPKSGTSLLEVVYLVSQAD 73
      : : | | | | | | | | | | | | | | | | | | | |
Db 7 EYEVFGFGRGLMDKRFYKYWEDVENFLARPDDLVIATPKSGTTWISVVYMIYKEGD 66

Qy 74 PDEIGLMNIDQLPVLEYPOP-----GLDI IKELTSPRLIKSHLPYRFLPSDLHNGDSKVI 129
      : : | | | | | | | | | | | | | | | | | | | |
Db 67 VEKCKEDAIFNRIPYLECRNEDLINGIKQLKEKESPRIVKTHLPKVLPA SFWEKNCCKMI 126

Qy 130 YMARNPDLVSYVYQFHRSLRTMSYRGTFQFCFRFMNDKLGYSWFEHVSQFWEHRMDS 189
      : : | | | | | | | | | | | | | | | | | | | |
Db 127 YLCRNADKDVAVSYVYFLMLMITSYPNPKSFSEFVEKFMQGVPGSWYDHWKAWWEKSKNS 186

Qy 190 NVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCH-----QLVDQ 236
      : : | | | | | | | | | | | | | | | | | | | |
Db 187 RVLFMFYEDMKEDIRREVVKLIEFLERKPSAELVDRIIQTSTFOEMKNPSTNYTMPEE 246

Qy 237 CCNAEALP-VGRGRVGLWKDIFTVSNNEKFDLVYKQMGKCDLTF 280
      : : | | | | | | | | | | | | | | | | | | | |
Db 247 MNQKVSFPMRKGIIGDNKNHFPPEALRERFDEHYKQMKDCTVKF 291
```



```
; PRIORITY FILING DATE: 2001-02-07
; PRIORITY APPLICATION NUMBER: 60/266,975
; PRIORITY FILING DATE: 2001-02-07
; PRIORITY APPLICATION NUMBER: 60/267,459
; PRIORITY FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 585
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Gallus gallus
US-10-072-012-585

Query Match      28.0%; Score 428; DB 4; Length 307;
Best Local Similarity 32.6%; Pred. No. 4.8e-34;
Matches 87; Conservative 61; Mismatches 97; Indels 22; Gaps 2;

Qy 36 EEIANFPVRPSDVWITVPKSGTSLLOEVVYLVSGQADPDEIGLMNIDEQLPVLE- 94
Db 37 DQVWNFKARPDDLAVATYAKAGTTWTOEIVDMIOQNGDIEKCRRASTYKRHPFLEWYIPD 96
Qy 95 -----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGSKVIYMARNPCKDLVSVYQFHR 147
Db 97 SSPGYSGLKUAEAMPSPRTKTHLPVQLVPPSPFWEQNKIIYVARNADNLVSYIHFHR 156
Qy 148 SLRTMSYRGTFQFCRRFMNDKLGYSWFEHVQEFWEHRMDSNVILFLKYEDMHRDLVTMV 207
Db 157 MNKVLDPDGTIEEFTKFMNGEVLGWSYDVHVGWAKADKRIILYFYEDMKENPKREI 216
Qy 208 EQLARFLVSGCDKAOQ-----EALTEHCHQLVDDQCCNAEALPVGRGVGLW 253
Db 217 QKIMFKLEKLDDEEVLNKKIYNTSPKIMKDNPMNTYTKDFVGVMDHSVSPFMRKSGVGDW 276
Qy 254 KDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 277 KNYFTVALNKKPQDQYKKKMDATSLVF 303

RESULT 28
US-10-199-330-5
; Sequence 5, Application US/10199330
; Publication No. US20020182681A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-5

Query Match      27.6%; Score 422; DB 4; Length 304;
Best Local Similarity 32.3%; Pred. No. 1.9e-33;
Matches 97; Conservative 59; Mismatches 124; Indels 20; Gaps 5;

Qy 1 MAESEAETPSTGGEFE-SKYFEFHGVRLPPFCRGHMEIANFPVRPSDVWITVPKSGTS 59
Db 1 MAKIEKNAPTWEKPELFNMEVDGVPPTLIILSKWEKVCNFOAKPDDLIATYPKSGTT 60
Qy 60 LLOEVVYLVSGQADPDEIGLMNIDEQLPVLE--YP---QPGLDIIKELTSPRLIKSHLPY 114
Db 61 WMHEILDMLNDGVEKCKRAQTLDRHAFLELKPHEKPDLEFVLENSSQLIKTHLPS 120
Qy 115 RFLPSDLHNGDSKVIYMARNPCKDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYS 174
Db 121 HLIPPSIWKENCKIYVARNPKDCLVSYIHFHRMASFMPDPQNLEEFYEKFMKSGVVGGS 180
Qy 175 WFEHVQEFWEHRMDSNVILFLKYEDMHRDLVTMVQELARFLVSGCDKAOLEALTEHCHQLV 234
Db 181 WFDHVQGWMAAKDMHRILYFYEDIKKDKREIKLFLKLEKDISEEILNKIIYHTSFDV 240
Qy 235 ---DQCCNAEALPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 MKQNPMTNYTTLPTSIMDHSISPFMRKMGPDGKNYFTVAQNEEFBDKDYQKMGAGTSLTF 300

RESULT 29
US-10-199-334-5
; Sequence 5, Application US/10199334
; Publication No. US20030008354A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-5

Query Match      27.6%; Score 422; DB 4; Length 304;
Best Local Similarity 32.3%; Pred. No. 1.9e-33;
Matches 97; Conservative 59; Mismatches 124; Indels 20; Gaps 5;

Qy 1 MAESEAETPSTGGEFE-SKYFEFHGVRLPPFCRGHMEIANFPVRPSDVWITVPKSGTS 59
Db 1 MAKIEKNAPTWEKPELFNMEVDGVPPTLIILSKWEKVCNFOAKPDDLIATYPKSGTT 60
Qy 60 LLOEVVYLVSGQADPDEIGLMNIDEQLPVLE--YP---QPGLDIIKELTSPRLIKSHLPY 114
Db 61 WMHEILDMLNDGVEKCKRAQTLDRHAFLELKPHEKPDLEFVLENSSQLIKTHLPS 120
Qy 115 RFLPSDLHNGDSKVIYMARNPCKDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYS 174
Db 121 HLIPPSIWKENCKIYVARNPKDCLVSYIHFHRMASFMPDPQNLEEFYEKFMKSGVVGGS 180
Qy 175 WFEHVQEFWEHRMDSNVILFLKYEDMHRDLVTMVQELARFLVSGCDKAOLEALTEHCHQLV 234
Db 181 WFDHVQGWMAAKDMHRILYFYEDIKKDKREIKLFLKLEKDISEEILNKIIYHTSFDV 240
Qy 235 ---DQCCNAEALPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 MKQNPMTNYTTLPTSIMDHSISPFMRKMGPDGKNYFTVAQNEEFBDKDYQKMGAGTSLTF 300

RESULT 30
```

```
US-10-199-329-5
; Sequence 5, Application US/10199329
; Publication No. US20030166189A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-5

Query Match      27.6%; Score 422; DB 4; Length 304;
Best Local Similarity 32.3%; Pred. No. 1.9e-33;
Matches 97; Conservative 59; Mismatches 124; Indels 20; Gaps 5;

Qy 1 MAESEAEPTSPGGEF-SKYFEFHGVRLPPFCRCGMEEIANFPVRPSDVWIVTPKSGTS 59
Db 1 MAKIEKNAPTWEKKPELFNIMEVDGVPPTLILSKWEKVCNFQAKPDDLIATPKSGTT 60
Qy 60 LLEQVYLVISQAGADPDDEIGLNMNIDEQLPVLE--YP---QPGLDIIKELTSPRLIKSHLPY 114
Db 61 WMHEILDMILNDGVEKCKRAQTLDRHAFLELKPHEKPDLEFVLEMSSPQLIKTHLPS 120
Qy 115 RFLPSDLHNGDSKVIYMARNPDKLVSYQFHRSLRTMSYRGTFQBFRCRRFMDKLGYS 174
Db 121 HLIPPSIWKENCKIYVARNPKOCLVSYHFRMASFMPDPQNLEEFYEKFMKGVVGS 180
Qy 175 WFEHVQEFWEHRMDSNVFLKYEYDMHRDLVTMVEQLARFLGVCDCQAQLEALTEHCHLV 234
Db 181 WFDHVKGWAAKDMHRILYLFYEDIKKDPKREIEKILKFLKDKISEILNKIIYHTSPDV 240
Qy 235 ---DQCCNAEALPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 MKQNPMTNTYTTLPTSIIMDHSISPFMRKMGPDGKNKYFTVAQNEEFDKDYQKKMAGSTLTF 300

RESULT 31
US-11-108-875-5
; Sequence 5, Application US/11108875
; Publication No. US20050196802A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-4
; CURRENT APPLICATION NUMBER: US/11/108,875
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: 10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40

US-10-199-329-5
; Sequence 5, Application US/09898570
; Patent No. US20020123612A1
; GENERAL INFORMATION:
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: 15966-776CIP
; CURRENT APPLICATION NUMBER: US/09/898,570
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/198,293
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/198,645
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/210,809
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/199,476
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,025
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/224,610
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/200,024
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/199,880
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/218,591
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/271,814
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/215,855
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 09/839,446
; PRIOR FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
```





```
Qy 81 NIDEQLPVLEY--POP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
Db 61 NTYDRHPFIEWTLPSPLNSGLDLANKMPSRPTLKLHPVHMLPPSPFWKENSIIYVARNA 120
Qy 136 KDLVSYVYQFHRSLRTMSYRGTFQFCRRFNMNDKLGYSWFEHVOEFWEHRMDSNVFLK 195
Db 121 KDLVSYVYFHRMKNLPPDGLGTYIEQFKAGKVLGWSYDHWKGMWDVQDKQHRILYLF 180
Qy 196 YEDMHRDLVTMVEQLARFLGVSCDKAQLAEALTEHCHQLV---DOCCNAEALP----- 244
Db 181 YEDMKEDPKREIKKIAKFLKDISEVLNKIIYHTSFDVMKENPMANYTTLPSSIMDHSI 240
Qy 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 SPFMRKMGPDWKNYFTTVAQSEDFEDYRRKMAGSNITF 279

RESULT 38
US-10-199-329-8
; Sequence 8, Application US/10199329
; Publication No. US20030166189A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 09/609,816
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-8
Query Match 27.4%; Score 419.5; DB 4; Length 283;
Best Local Similarity 32.6%; Pred. No. 3e-33;
Matches 91; Conservative 57; Mismatches 112; Indels 19; Gaps 4;
Qy 21 EFHGVRLPPFCRGKMEETANFPVRPSDVWIVTYPKSGTSLLOEVVYLVSQGADPDDEIGLM 80
Db 1 EVNGILMSKLSMDNWDKIWNFQAKPDDLLIATYAKAGTTWTQEIYVDMIQNDGDVQKQORA 60
Qy 81 NIDEQLPVLEY--POP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
Db 61 NTYDRHPFIEWTLPSPLNSGLDLANKMPSRPTLKLHPVHMLPPSPFWKENSIIYVARNA 120
Qy 136 KDLVSYVYQFHRSLRTMSYRGTFQFCRRFNMNDKLGYSWFEHVOEFWEHRMDSNVFLK 195
Db 121 KDLVSYVYFHRMKNLPPDGLGTYIEQFKAGKVLGWSYDHWKGMWDVQDKQHRILYLF 180
Qy 196 YEDMHRDLVTMVEQLARFLGVSCDKAQLAEALTEHCHQLV---DOCCNAEALP----- 244
Db 181 YEDMKEDPKREIKKIAKFLKDISEVLNKIIYHTSFDVMKENPMANYTTLPSSIMDHSI 240
Qy 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 SPFMRKMGPDWKNYFTTVAQSEDFEDYRRKMAGSNITF 279

US-10-199-329-9
; Sequence 9, Application US/10199329
; Publication No. US20030166189A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 09/609,816
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-9
Query Match 27.4%; Score 419.5; DB 4; Length 283;
Best Local Similarity 32.6%; Pred. No. 3e-33;
Matches 91; Conservative 57; Mismatches 112; Indels 19; Gaps 4;
Qy 21 EFHGVRLPPFCRGKMEETANFPVRPSDVWIVTYPKSGTSLLOEVVYLVSQGADPDDEIGLM 80
Db 1 EVNGILMSKLSMDNWDKIWNFQAKPDDLLIATYAKAGTTWTQEIYVDMIQNDGDVQKQORA 60
Qy 81 NIDEQLPVLEY--POP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
Db 61 NTYDRHPFIEWTLPSPLNSGLDLANKMPSRPTLKLHPVHMLPPSPFWKENSIIYVARNA 120
Qy 136 KDLVSYVYQFHRSLRTMSYRGTFQFCRRFNMNDKLGYSWFEHVOEFWEHRMDSNVFLK 195
Db 121 KDLVSYVYFHRMKNLPPDGLGTYIEQFKAGKVLGWSYDHWKGMWDVQDKQHRILYLF 180
Qy 196 YEDMHRDLVTMVEQLARFLGVSCDKAQLAEALTEHCHQLV---DOCCNAEALP----- 244
Db 181 YEDMKEDPKREIKKIAKFLKDISEVLNKIIYHTSFDVMKENPMANYTTLPSSIMDHSI 240
Qy 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 SPFMRKMGPDWKNYFTTVAQSEDFEDYRRKMAGSNITF 279

RESULT 39
US-10-199-329-9
; Sequence 9, Application US/10199329
```

```
; Publication No. US20030166189A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 09/609,816
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-9
Query Match 27.4%; Score 419.5; DB 4; Length 283;
Best Local Similarity 32.6%; Pred. No. 3e-33;
Matches 91; Conservative 57; Mismatches 112; Indels 19; Gaps 4;
Qy 21 EFHGVRLPPFCRGKMEETANFPVRPSDVWIVTYPKSGTSLLOEVVYLVSQGADPDDEIGLM 80
Db 1 EVNGILMSKLSMDNWDKIWNFQAKPDDLLIATYAKAGTTWTQEIYVDMIQNDGDVQKQORA 60
Qy 81 NIDEQLPVLEY--POP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
Db 61 NTYDRHPFIEWTLPSPLNSGLDLANKMPSRPTLKLHPVHMLPPSPFWKENSIIYVARNA 120
Qy 136 KDLVSYVYQFHRSLRTMSYRGTFQFCRRFNMNDKLGYSWFEHVOEFWEHRMDSNVFLK 195
Db 121 KDLVSYVYFHRMKNLPPDGLGTYIEQFKAGKVLGWSYDHWKGMWDVQDKQHRILYLF 180
Qy 196 YEDMHRDLVTMVEQLARFLGVSCDKAQLAEALTEHCHQLV---DOCCNAEALP----- 244
Db 181 YEDMKEDPKREIKKIAKFLKDISEVLNKIIYHTSFDVMKENPMANYTTLPSSIMDHSI 240
Qy 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 SPFMRKMGPDWKNYFTTVAQSEDFEDYRRKMAGSNITF 279

RESULT 40
US-11-108-875-8
; Sequence 8, Application US/11108875
; Publication No. US20050196802A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669DIV-4
; CURRENT APPLICATION NUMBER: US/11/108,875
; PRIOR FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: 10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/192,408
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
```





```
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-10

Query Match      26.7%; Score 408.5; DB 4; Length 283;
Best Local Similarity 31.5%; Pred. No. 3.9e-32;
Matches 88; Conservative 57; Mismatches 115; Indels 19; Gaps 3;

QY 21 EPHGVRPLPPFCRGMKEEIANFVRPSDVWIVTPKSGTSLQEVVYLVSGQADPDPEIGLM 80
Db 1 EVNGILSKMMSSENWDKIWNFAKQKPDLLIATYAKAGTTWTQEIIVDMIQNDGVDQKQRA 60
QY 81 NIDGQLPVLEYPQP-----GLDIIKELTSPRIKSHLPYRELPSDLHNGDSKVIYMARNP 135
Db 61 NYDRHPFIEWTLPPPLNSGLDLANKMPSPTLTKHLPVQMLPSPFWKENSQIIIVARNA 120
QY 136 KDLVSVYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDSNVLEFLK 195
Db 121 KDLVSVYYFSSRMNKMPLDPGTLGEYIETFKAGKVLGWSWDYHVKGMWDVKDKHRLILYLF 180
QY 196 YEDMHRDLVTWVEQLARFLGVSCKAQLEALTEHCHQLV---DOCCNAEALP----- 244
Db 181 YEDMKEDPKREIKKIVKFLKDISSEVLNKIIHTSFDMKQNPANVTTLTLPSSIMDSHI 240
QY 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 SPFMRKGMFGDKWKNYFTVAQSEDFEDYRKQWAGSTITF 279

RESULT 47
US-10-072-012-584
; Sequence 584, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072.012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 584
; LENGTH: 304
```







Db 165 YYHYFKLLHGMN--GDFEQFVDFLEGHTPMGSYWRHVLFPWKRQDDNVLFKDYEDMVK 222

Qy 202 DLVTWVEQIARPLGVS--CDKAQLEALTEHC---HOLVDQCCNAE-ALP-----VGRGR 249

Db 223 DLPSVVRRCARGLVQSLDVLSTLQKLDHLTFDKWRANKAVNLEKLLPSSSKFIRNGK 282

Qy 250 VGLWKDIFTVSNMKEFDLVYKQKMGKCDLTFDF 282

Db 283 IGDWRNHNMGNSERFDEWTEKMRGSLNFDY 315

RESULT 53

US-10-199-330-4

Sequence 4, Application US/10199330

Publication No. US20020182681A1

GENERAL INFORMATION:

APPLICANT: WOODAGE, Trevor et al.

TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING

TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN

TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF

FILE REFERENCE: CL000669DIV-1

CURRENT APPLICATION NUMBER: US/10/199,330

PRIOR FILING DATE: 2002-07-22

PRIOR APPLICATION NUMBER: 09/609,816

PRIOR FILING DATE: 2001-03-27

PRIOR APPLICATION NUMBER: 60/192,408

PRIOR FILING DATE: 2000-03-27

PRIOR APPLICATION NUMBER: 60/212,725

PRIOR FILING DATE: 2000-06-20

PRIOR APPLICATION NUMBER: 09/609,816

PRIOR FILING DATE: 2000-07-03

NUMBER OF SEQ ID NOS: 23

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 4

LENGTH: 304

TYPE: PRT

ORGANISM: Homo sapiens

US-10-199-330-4

Query Match 26.2%; Score 401; DB 4; Length 304;

Best Local Similarity 31.0%; Pred. No. 2.4e-31;

Matches 93; Conservative 62; Mismatches 125; Indels 20; Gaps 5;

Qy 1 MAESEAEPTSPGFE-SKYFEFHGVRLPPFCRCGMEEIANFPVRPSDVWIVTPKSGTS 59

Db 1 MAKIEKNAPTMEKKPELFNMEVDGVTLLLSKEWKEKVPQAKPDDLILATPKSGTT 60

Qy 60 LLEQVVYLVSGADPDEIGLWNIDEQLPVLE--YP--QPGLDIIKELTSRLIKSHLPY 114

Db 61 WMHEILDMLNDGDVEKCKRAQTLDRHAFLEKLPKHEKPDLEFVLEMSSPOLIKTHLPS 120

Qy 115 RFLPSDLHNGDSKVIYMARNPDLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYS 174

Db 121 HLIPPSIWKENCKIVYVARNPKDCLVSYHFRMASFMPDQNLEEFYEFKEMSGKVVGGS 180

Qy 175 WFEHVQEFWEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCKQAQLEALTEHCOLV 234

Db 181 WFDHVKGWAAKDMHRILYLFYEDIKKNPKHEIHKLVLEFLEKTSWGDVINKIVHHTSPDV 240

Qy 235 ---DQCCNAEALP-----VGRGRVGLWKDIFTVSNMKEFDLVYKQKMGKCDLTF 280

Db 241 MKDNPMAHTAVPAHIFNHSISKFMKGMPCGMDKWNHFTVALNENFDKHYEKKMAGSTLNF 300

RESULT 54

US-10-199-334-4

Sequence 4, Application US/10199334

Publication No. US20030008354A1

GENERAL INFORMATION:

APPLICANT: WOODAGE, Trevor et al.

TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING

TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN

TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF

FILE REFERENCE: CL000669DIV-3

CURRENT APPLICATION NUMBER: US/10/199,334

PRIOR FILING DATE: 2002-07-22

PRIOR APPLICATION NUMBER: 09/609,816

PRIOR FILING DATE: 2001-03-27

PRIOR APPLICATION NUMBER: 60/192,408

PRIOR FILING DATE: 2000-03-27

PRIOR APPLICATION NUMBER: 60/212,725

PRIOR FILING DATE: 2000-06-20

PRIOR APPLICATION NUMBER: 09/609,816

PRIOR FILING DATE: 2000-07-03

NUMBER OF SEQ ID NOS: 23

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 4

LENGTH: 304

TYPE: PRT

ORGANISM: Homo sapiens

US-10-199-334-4

Query Match 26.2%; Score 401; DB 4; Length 304;

Best Local Similarity 31.0%; Pred. No. 2.4e-31;

Matches 93; Conservative 62; Mismatches 125; Indels 20; Gaps 5;

Qy 1 MAESEAEPTSPGFE-SKYFEFHGVRLPPFCRCGMEEIANFPVRPSDVWIVTPKSGTS 59

Db 1 MAKIEKNAPTMEKKPELFNMEVDGVTLLLSKEWKEKVPQAKPDDLILATPKSGTT 60

Qy 60 LLEQVVYLVSGADPDEIGLWNIDEQLPVLE--YP--QPGLDIIKELTSRLIKSHLPY 114

Db 61 WMHEILDMLNDGDVEKCKRAQTLDRHAFLEKLPKHEKPDLEFVLEMSSPOLIKTHLPS 120

Qy 115 RFLPSDLHNGDSKVIYMARNPDLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYS 174

Db 121 HLIPPSIWKENCKIVYVARNPKDCLVSYHFRMASFMPDQNLEEFYEFKEMSGKVVGGS 180

Qy 175 WFEHVQEFWEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCKQAQLEALTEHCOLV 234

Db 181 WFDHVKGWAAKDMHRILYLFYEDIKKNPKHEIHKLVLEFLEKTSWGDVINKIVHHTSPDV 240

Qy 235 ---DQCCNAEALP-----VGRGRVGLWKDIFTVSNMKEFDLVYKQKMGKCDLTF 280

Db 241 MKDNPMAHTAVPAHIFNHSISKFMKGMPCGMDKWNHFTVALNENFDKHYEKKMAGSTLNF 300

RESULT 55

US-10-199-329-4

Sequence 4, Application US/10199329

Publication No. US20030166189A1

GENERAL INFORMATION:

APPLICANT: WOODAGE, Trevor et al.

TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING

TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN

TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF

FILE REFERENCE: CL000669DIV-2

CURRENT APPLICATION NUMBER: US/10/199,329

PRIOR FILING DATE: 2002-07-22

PRIOR APPLICATION NUMBER: 09/609,816

PRIOR FILING DATE: 2001-03-27

PRIOR APPLICATION NUMBER: 60/192,408

PRIOR FILING DATE: 2000-03-27

PRIOR APPLICATION NUMBER: 60/212,725

PRIOR FILING DATE: 2000-06-20

PRIOR APPLICATION NUMBER: 09/609,816

PRIOR FILING DATE: 2000-07-03

NUMBER OF SEQ ID NOS: 23

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 4

LENGTH: 304

TYPE: PRT

ORGANISM: Homo sapiens

US-10-199-329-4

Query Match 26.2%; Score 401; DB 4; Length 304;





Db 241 MKDNPMANHTAVPAHFNHSISKFMKRGMPGDWKNHFTVAMNENFDKHYEKKMAGSTLNF 300

RESULT 60

US-10-072-012-206  
; Sequence 206, Application US/10072012  
; Publication No. US20040033493A1

GENERAL INFORMATION:

APPLICANT: Tchernev, Velizar

APPLICANT: Spytek, Kimberly

APPLICANT: Zerhusen, Bryan

APPLICANT: Patturajan, Meera

APPLICANT: Shimkets, Richard

APPLICANT: Li, Li

APPLICANT: Gangolli, Esha

APPLICANT: Padigaru, Muralidhara

APPLICANT: Anderson, David W.

APPLICANT: Rastelli, Luca

APPLICANT: Miller, Charles E.

APPLICANT: Gerlach, Valerie

APPLICANT: Taupier Jr, Raymond J.

APPLICANT: Gusev, Vladimir Y.

APPLICANT: Colman, Steven D.

APPLICANT: Wolenc, Adam R.

APPLICANT: Pena, Carol E. A

APPLICANT: Furtak, Katarzyna

APPLICANT: Groesse, William M.

APPLICANT: Alsobrook II, John P.

APPLICANT: Lepley, Denise M.

APPLICANT: Rieger, Daniel K.

APPLICANT: Burgess, Catherine E.

;; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

;; FILE REFERENCE: 21402-258

;; CURRENT APPLICATION NUMBER: US/10/072,012

;; CURRENT FILING DATE: 2002-01-31

;; PRIOR APPLICATION NUMBER: 60/265,102

;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: 60/265,514

;; PRIOR FILING DATE: 2001-01-31

;; PRIOR APPLICATION NUMBER: 60/265,517

;; PRIOR FILING DATE: 2001-01-31

;; PRIOR APPLICATION NUMBER: 60/265,412

;; PRIOR FILING DATE: 2001-01-31

;; PRIOR APPLICATION NUMBER: 60/265,395

;; PRIOR FILING DATE: 2001-01-31

;; PRIOR APPLICATION NUMBER: 60/266,406

;; PRIOR FILING DATE: 2001-02-02

;; PRIOR APPLICATION NUMBER: 60/266,767

;; PRIOR FILING DATE: 2001-02-05

;; PRIOR APPLICATION NUMBER: 60/267,057

;; PRIOR FILING DATE: 2001-02-07

;; PRIOR APPLICATION NUMBER: 60/266,975

;; PRIOR FILING DATE: 2001-02-07

;; PRIOR APPLICATION NUMBER: 60/267,459

;; PRIOR FILING DATE: 2001-02-08

;; Remaining Prior Application data removed - See File Wrapper or PALM.

;; NUMBER OF SEQ ID NOS: 1391

;; SOFTWARE: PatentIn Ver. 2.1

;; SEQ ID NO 206

;; LENGTH: 295

;; TYPE: PRT

;; ORGANISM: Homo sapiens

;; OTHER INFORMATION: 206

Query Match 26.0%; Score 397.5; DB 4; Length 295;

Best Local Similarity 32.2%; Pred. No. 5.2e-31;

Matches 85; Conservative 58; Mismatches 102; Indels 19; Gaps 4;

QY 36 EEIANFVRPSDWMVTVTPKSGTSLQEVVYLVSGADPDEIGLMNIDEQPVLE--YP- 92

Db 28 EKVCNFOAKPDOLIATPKSGTTPWMEHILDMILNDGVEKCKRAQTLDRAHFLKLPFH 87

QY 93 --QPGLDIIEKLTSPRIKSHLPYRFLPSDLHNGDSKVIYARNPKDLVWSYQPHRSLR 150

Db 88 KEKPDLEFVLEMSSPQIKTHLPShLIPSPSTWKNCKIVYVARNPKDCLVSYVHFRNAS 147

QY 151 TMSYRGTFQEFRCRPMNDKLGYSWFHGVQFWEHRMDSNVLFKLYEDMHRDLVTMVEQL 210

Db 148 FMPDPQNLLEEFYKFMGSKVVGSGWFDHMKMGWAAKQMHRIYLFYEDIKKNPKHEIHKV 207

QY 211 ARFLGVSCDKAQLEALTEHCHQLV---DQCNAALP-----VGRGRVGLWKDI 256

Db 208 LEFLEKTWSGDVINKIVHHTSFDYMKONPMANHTAVPAHFNHSISKFMKRGMPGDWKNH 267

QY 257 FTVMNEKFDLVYKQKMGKCDLTF 280

Db 268 FTVAMNENFDKHYEKKMAGSTLNF 291

RESULT 61

US-09-898-570-26

;; Sequence 26, Application US/09898570

;; Patent No. US20020123612A1

GENERAL INFORMATION:

APPLICANT: GERLACH, VALERIE L.

APPLICANT: ELLERMAN, KAREN

APPLICANT: MACDOUGALL, JOHN R.

APPLICANT: SMITHSON, GLENNDA

;; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND

;; FILE REFERENCE: 15966-776CIP

;; CURRENT APPLICATION NUMBER: US/09/898,570

;; CURRENT FILING DATE: 2001-07-03

;; PRIOR APPLICATION NUMBER: 60/198,293

;; PRIOR FILING DATE: 2000-04-19

;; PRIOR APPLICATION NUMBER: 60/198,645

;; PRIOR FILING DATE: 2000-04-20

;; PRIOR APPLICATION NUMBER: 60/210,809

;; PRIOR FILING DATE: 2000-06-09

;; PRIOR APPLICATION NUMBER: 60/199,476

;; PRIOR FILING DATE: 2000-04-26

;; PRIOR APPLICATION NUMBER: 60/200,025

;; PRIOR FILING DATE: 2000-04-26

;; PRIOR APPLICATION NUMBER: 60/224,610

;; PRIOR FILING DATE: 2000-08-11

;; PRIOR APPLICATION NUMBER: 60/200,024

;; PRIOR FILING DATE: 2000-04-26

;; PRIOR APPLICATION NUMBER: 60/199,880

;; PRIOR FILING DATE: 2000-04-26

;; PRIOR APPLICATION NUMBER: 60/218,591

;; PRIOR FILING DATE: 2000-07-17

;; PRIOR APPLICATION NUMBER: 60/271,814

;; PRIOR FILING DATE: 2001-02-27

;; PRIOR APPLICATION NUMBER: 60/215,855

;; PRIOR FILING DATE: 2000-07-03

;; PRIOR APPLICATION NUMBER: 09/839,446

;; PRIOR FILING DATE: 2001-04-19

;; NUMBER OF SEQ ID NOS: 58

;; SOFTWARE: PatentIn Ver. 2.1

;; SEQ ID NO 26

;; LENGTH: 305

;; TYPE: PRT

;; ORGANISM: Unknown Organism

;; FEATURE:

;; OTHER INFORMATION: Description of Unknown Organism: POLYX

;; OTHER INFORMATION: h\_nh0443k08\_A

US-09-898-570-26

Query Match 25.9%; Score 396.5; DB 3; Length 305;

Best Local Similarity 30.9%; Pred. No. 6.9e-31;

Matches 93; Conservative 60; Mismatches 127; Indels 21; Gaps 5;

QY 1 MAESEAETPPTGSGFE-SKYFEFHGVLPPFCRGMBEIANFPVRPSDWMVTVTPKSGTS 59

Db 1 MAKIEKNAPTMEKPELFNIMEVDGVPTLILSKWEKVCNFOAKPDOLIATYTPKSGTT 60

Qy	60	LLQEVVYLVSQGADPDEIGLMMIDQLPVLRE--YP---OPGLDIKELTSPLRIKSHLPY	114
Db	61	WMHEILDMILNDGDEVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSPOLIKTHPS	120
Qy	115	RFLPSDLHGDSKVYIMARNPKDVLVSYQYFHRSLRTMSYRGTFQEFRCRRWMDKLGYS	174
Db	121	HLIIPSIWKENCKIVYVAARNPKDCLVSYVYFHRMASFMPDPQNLSEFEYFKFMSGKEFGS	180
Qy	175	WFEHVQEFWEHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTETGHQOLV	234
Db	181	WFDHVYKGWAAKDMHRILYLFYEDIKQNPKEIHKVLEFLEKTSWGDVINKIVVHSTFVD	240
Qy	235	---DOCCNAEALPV-----GGRVGLWKDIFTVSMNEKFDLVYVKOMGKCDLT	279
Db	241	MKDNPMAHNTAVPAHI FNHSISKFMKGGMPGDWKNHFTVALNEAFDKHYEKKMGAGSTLN	300
Qy	280	F 280	
Db	301	F 301	

RESULT 62

```

US-09-839-446-26
; Sequence 26, Application US/09839446
; Publication No. US20030050232A1
; GENERAL INFORMATION:
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN
; METHOD OF INVENTION: METHODS OF USING
; FILE REFERENCE: 15966-776
; CURRENT APPLICATION NUMBER: US/09/839-446-26
; CURRENT FILING DATE: 2003-04-19
; PRIOR APPLICATION NUMBER: 60/198,299
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/198,643
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/210,803
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/199,478
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,022
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/224,611
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/200,022
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/199,888
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/218,598
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/271,811
; PRIOR FILING DATE: 2001-02-27
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of t
; OTHER INFORMATION: h_rh0443k08_A
US-09-839-446-26

```

Query Match 25.9%; Score 396.5; DB 3; Length 305;

Best Local Similarity 30.9%; Pred. No. 6.9e-31;

Matches	93	Conservative	60	Mismatches	127	Indels	21	Gaps	5
---------	----	--------------	----	------------	-----	--------	----	------	---

QY 1 MAESEAETPSTPGEFE-SKYFEFHGVRLPFFCRGKMEEIANFPVRPSDVWIVTYPKSGTS 59

Db 1 MAK1EKNAPTWEKKKPLFN1NEVDGVTLLSKEWKEKVCNFOAKPDDLLATVPKSGTT 60

Query Match 25.2%: Score 385.5: DB 4: Length 297:

Qy	60	LLQEWVYVLSQAGDPDEIGLWNIDQLPVLE--YP---OPGLDIIKELTSPLIKSHLPY	114
Db	61	WMHEILDMILNDGVEKCKRAQTLLDRHAFLEKFPHEKPKDLEFVLEMSSPQIILKTHLPS	120
Qy	115	RFLPSDLHNGDSKVIYVARNPKDLVSYQYFHRSLRTMSYRGTFQFCRCRRFNMNDKLGVS	174
Db	121	HLIPIISWKENCKIVVYARNPKDCLVSYHFRMASFMPDQNLSEEFYKCFMSGKGFGS	180
Qy	175	WFEHVQEFWEHRMDSNVFLKYEDMHRDLVTWVQLARFLVGSCKAQLEALTEHCHQLV	234
Db	181	WFDHVKGWAAKQMHRIYLIFYEDIKQPKHEIHKVLEFLEKTSWGDVINKIVHHTSFVD	240
Qy	235	---DQCCNAEALPV-----GGRVGLWKDIFTVSNNKRFDLVYKQKMGKCDLT	279
Db	241	MKNPMMNHTAVPAHIFNHSISKPMKCGMGPDGWNKHTVALNENFDKHYEKKMGAGSTLN	300
Qy	280	F 280	
Db	301	F 301	

RESULT 63

```

US-10-381-898-4
; Sequence 4, Application US/10381898
; Publication NO. US20040086887A1
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Valda; BAUGHN, Mariah R.;
; APPLICANT: BOROWSKY, Mark L.; DING, Li;
; APPLICANT: DUGGAN, Brendan; ELLIOTT, Vicki S
; APPLICANT: GANDHI, Ameena R.; GRIFFIN, Jenni
; APPLICANT: HAFALIA, April J.A.; ISON, Craig G.;
; APPLICANT: KHAN, Farrah A.; LAL, Preeti G.;
; APPLICANT: LEE, Ernestine A.; LU, Dyung Aina
; APPLICANT: ARVIZO, Chandra S.; POLICKY, Jenni
; APPLICANT: RAMKUMAR, Jayalaxmi; RING, Huijun
; APPLICANT: SANGUANWALA, Madhusudan M.;
; APPLICANT: TANG, Y. Tom; TRIBOULEY, Catherine
; APPLICANT: CHAWLA, Narinder K.; WALSH, Roger
; APPLICANT: WARREN, Bridget; XU, Yuming;
; APPLICANT: YANG, Junming; YAO, Monique; YUE,
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0233 USN
; CURRENT APPLICATION NUMBER: US/10/381.898
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/US01/30662
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/236,947
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/238,864
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/242,323
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/247,581
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/249,519
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252,834
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/250,567
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 71680316CD1
US-10-381-898-4

```

### Query Match

```
Best Local Similarity 32.4%; Pred. No. 8.4e-30;
Matches 84; Conservative 55; Mismatches 105; Indels 15; Gaps 3;

Qy 36 BEIANFPVRSDVWIVTPKSGTSLQVYVLSQGADPDEIGLNMNIDEQLPVLEYPOPG 95
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 36 DKINWFQAKPDLLATTPYKAGTTWMHEILDMLNDGVEKCKRAQTLDRAHAFLELKEP- 94
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 96 LDIIKELTSPRLIKSHLYRFLPSDLHNGDSKVIYMAENPKDLVSYVYQFHRSLRTMSYR 155
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 95 LEFVLEMSPOLIKTHLPSHILPPSIWKENCKIVVYARNPKDCLVSYVYHFRMASFMPDP 154
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 156 GTFOFCRRFNMNDKLGYSWFHEVQFWEHRMDSNVLFKYEDMHRDLVTWVQOLARFLG 215
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 155 QNLSEFYEKFMGKVGKVSFHDVKGWAAKDTHRILYLFYEDIKKPKHEIHKVLEFL 214
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 216 VSCDKAQLAEALTEHCHQLV---DQCNABALP-----VGRGRVGLWKDIFTVSM 261
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 215 KTLSGDVINKIVHHTSFDMKNPMAHTAVPAHIFNHSISKFMKGMGPDGKNHFTVAM 274
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 262 NEKEDLVYKQMGKCDLTF 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 275 NENFDKHYEKKMAGSTLNF 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 64
US-10-072-012-204
; Sequence 204, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zethusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shankets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pens, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975

; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 204
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-204

Query Match 25.0%; Score 383; DB 4; Length 324;
Best Local Similarity 28.8%; Pred. No. 1.7e-29;
Matches 92; Conservative 62; Mismatches 126; Indels 40; Gaps 6;

Qy 1 MAESEASTPTPGFE-SKYFEFHGVRLLPPPCRGKMBEIANFPVRPSDV-----48
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MAKIEKNAPTMEKKPELNFMEVDGVTLLLSKEWKEKVCNFOAKPDDLIIHASMLYLTG 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 49 -----WIVTPYKSGTSLQVYVLSQGADPDEIGLNMNIDEQLPVLE--YP---QP 94
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 KLPEEDHQAWLGNYPKSGTTMMHEILDMLNDGVEKCKRAQTLDRAHAFLELKEPHEKP 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 95 GLDIIKELTSPRLIKSHLYRFLPSDLHNGDSKVIYMAENPKDLVSYVYQFHRSLRTMSY 154
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 DLEFVLEMSPOLIKTHLPSHILPPSIWKENCKIVVYARNPKDCLVSYVYHFRMASFMPD 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 155 RGTFQFCRRFNMNDKLGYSWFHEVQFWEHRMDSNVLFKYEDMHRDLVTWVQOLARFL 214
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 PQLNEEFYEKFMGKVGKVSFHDVKGWAAKDTHRILYLFYEDIKKPKHEIHKVLEFL 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 215 VSCDKAQLAEALTEHCHQLV---DQCNABALP-----VGRGRVGLWKDIFTVSM 260
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 EKTWSGDVINKIVHHTSFDMKNPMAHTAVPAHIFNHSISKFMKGMGPDGKNHFTVA 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 261 MNEKFDLVYKQMGKCDLTF 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 NENFDKHYEKKMAGSTLNF 320
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 65
US-09-989-442-149
; Sequence 149, Application US/09989442
; Publication No. US20030013649A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0208
; CURRENT APPLICATION NUMBER: US/09/989,442
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
```



; PRIOR APPLICATION NUMBER: 60/216,647  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,267  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/216,880  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,270  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/251,869  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/235,834  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/234,274  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/234,223  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/228,924  
; PRIOR FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/224,518  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,369  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/224,519  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,964  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/241,809  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/249,299  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/236,327  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/241,785  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/244,617  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 60/225,268  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,368  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/251,856  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/229,344  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/234,997  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: 60/229,343  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,345  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,287  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,513  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/231,413  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/229,509  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/236,367  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/237,039  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,038  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/236,370  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/236,802  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,037  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,040

; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/240,960  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/239,935  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/239,937  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/241,787  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/246,474  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/246,532  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/249,216  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,210  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/226,681  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,759  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/225,213  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/227,182  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,214  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/235,836  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/230,438  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/215,135  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 60/225,266  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/249,218  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,208  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,213  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,212  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,207  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,245  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,244  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,217  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,211  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,215  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,264  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,214  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,297  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/232,400  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/231,242  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/232,081  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/232,080  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,414  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,244  
; PRIOR FILING DATE: 2000-09-08

```

; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,065
; PRIOR FILING DATE: 2000-09-14

Query Match      25.0%; Score 382; DB 3; Length 76;
Best Local Similarity 95.9%; Pred. No. 2.8e-30; Indels 0; Gaps 0;
Matches 71; Conservative 1; Mismatches 2;

Qy 90 EYPOGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYARNPKDLVVSYYQFHRSL 149
Db 1 KYPOGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYARNPKDLVVSYYQFHRSL 60

Qy 150 RTMSYRGTFQBFQR 163
Db 61 RTMSYRGXFQBFQR 74

RESULT 66
US-10-756-149-5061
; Sequence 5061, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 5061
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-5061

Query Match      24.9%; Score 381; DB 5; Length 350;
Best Local Similarity 33.3%; Pred. No. 3e-29;
Matches 94; Conservative 52; Mismatches 108; Indels 28; Gaps 7;

Qy 12 PGFESKYFEHGVRLPFCRG--KMEIA----NFPVRPSDVMIVTPKSGTSLLOEVV 65
Db 13 PGE----YFRYKGV---PFPVGLYSLSISLAENTQDVRDDDIIFIITPKSGTWTWIEII 65
Qy 66 YLVSQGADPDIGLNMNIDEQLPVLEYPOPGLDIIKELTSPRLIKSHLPYRFLPSDLHNGD 125
Db 66 CLIIKEGDPWSIRSPWIWEPACE-TIVGAFSLPDQYSPRLMSHLDIQTIFKAFSSK 124
Qy 126 SKVIYARNPKDLVVSYYQFHRSLRTSYRGTFQBFCCRFRMNDKLGYSWFHVOEFWEH 185
Db 125 AKVIYMGNRPRDVSLSYHYSKIAGQLKDPGTPOQLRDFLKGVEQFGSWFDHIKGLMRM 184
```

```

Qy 186 RMDSNVLPKXYEDMHRDLVTMVEQLARFLGVSCDKAOLEALTEH---CHQLVDOCCNAEA 242
Db 185 KGDNFLFITEEELQQLQSGVERICGFLGPLGKEALGVSVAHSTFSAMKANTMSNYTL 244
Qy 243 LP-----VGRGRVGLWKDIFTVSMNEKFDLVYKQKM 273
Db 245 LPPSLDHRRGAFRLKGVCGDKWNHFTVAQSEAFDRAYRQKM 286
```

## RESULT 67

```

US-11-097-143-11343
; Sequence 11343, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11343
; LENGTH: 338
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-11343
```

```

Query Match      23.6%; Score 360.5; DB 6; Length 338;
Best Local Similarity 31.0%; Pred. No. 3.3e-27;
Matches 89; Conservative 51; Mismatches 112; Indels 35; Gaps 7;
```

```

Qy 9 PSTP--GEFESKYFEHGVRLPFCRCGMEEIANFPVRPSDVMIVTPKSGTSLLOEVV 66
Db 34 PSVPVGVNWEQRF-----CRLADTFQPLDRVYDFEVRDDDDVWIVTLPCGTTWMQELAW 88
Qy 67 LVSGADPDIGLNMNIDEQLPVLEY-----PQPGLDIIKELTSPRLIKSHLPYRFLPS 119
Db 89 LVINECDFETAKSVDLTHRSFPFLEFNGVNVNPHDTTAAANALSPRLIKSHLPANMLPR 148
Qy 120 DLHNGDSKVIYARNPKDLVVSYYQFHRSLRTSYRGTFQBFCCRFRMNDKLGYSWFH 179
Db 149 QIWSKRPKIIVVYRNPKDAAISYFHHWRGM--VGYYQTQKDFMHSFIDGVYVNFPCWPHI 206
Qy 180 QEFWEHRMDSNVLPKYEDMHRDLVTMVEQLARFLGVSCDKAOLEALTEHC--HQLVDO- 236
Db 207 LDFMQLRHEPNIFTSYRMRKQGLQGVISEVAQFLERSVSEQEQMQQRHLSFESMRDNP 266
Qy 237 CCN-----AALPVGRGVGLMKDIFTVSMNEKFDL 267
Db 267 ACHVKEFESKMAAGREVEEFRRVRRGVGVGSHKDELTAIIRFDL 313
```

## RESULT 68

```
US-11-097-143-14628
; Sequence 14628, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14628
; LENGTH: 346
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-14628

Query Match      23.5%; Score 359.5; DB 6; Length 346;
Best Local Similarity 32.3%; Pred. No. 4.2e-27;
Matches 82; Conservative 35; Mismatches 88; Indels 49; Gaps 5;

QY 22 FHGVR-----LPPFCRGKMEETANFVRPSDVWIVTPKSGTSLQEVVLYVS 69
Db 27 FGEITGTVQVSGEYFFPHKYKDEARYNFPEARDDVWIATVRSGITWTQELIWLVA 86

QY 70 QGADPDEIGLNMIDQLPVLEYP-----QPGLDIK 100
Db 87 NGLDFEHAQERPLTERFPFPEFPLFVHPKIKEELOENRDSABALEFIEKIARPGVEALS 146

QY 101 EL--TSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDKLVVSYQFHRSLRTMSYRGTF 158
Db 147 EIPRSQRRIKTHFPFSLMPPSVLEKKKXVIYVRDPKDVAVSYVHLNRLFRTOGVGDF 206

QY 159 QSECFRRFMDKLGYSW---FEHVQFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLG 215
Db 207 ERYWYFQV---GLNPWLPYSHVKEAREHAHLSNVFLURYEDMLADLPAINSTIASFLPE 263

QY 216 VSCDKAQLAEALTEH 229
Db 264 CPPKPEDMDRLLDH 277

RESULT 69
US-11-152-635-11
; Sequence 11, Application US/11152635
; Publication No. US20060024708A1
; GENERAL INFORMATION:
; APPLICANT: Sinclair, James
; APPLICANT: Sinclair, Philip
; TITLE OF INVENTION: Porcine Sulfotransferase 2A1 Polynucleotide Sequence, Protein, an
; TITLE OF INVENTION: Methods of Use for Same
; FILE REFERENCE: P06815US01
; CURRENT APPLICATION NUMBER: US/11/152,635
; CURRENT FILING DATE: 2005-06-14
```

```
; PRIOR APPLICATION NUMBER: 60/580,540
; PRIOR FILING DATE: 2005-06-17
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-152-635-11

Query Match      23.3%; Score 356.5; DB 6; Length 285;
Best Local Similarity 31.0%; Pred. No. 6.5e-27;
Matches 83; Conservative 59; Mismatches 109; Indels 17; Gaps 5;

QY 22 FHGVRLLPP--FCRGKMEIAN-FVVRPSDVWIVTPKSGTSLQEVVLYVSGADPDEIG 78
Db 8 FEGIAFPTMGFRSETLRKVRDEFVIRDEVDVILTYPKSGTNWLABILCLMHKSGDAKWIQ 67

QY 79 LMNIDEQLPVLEYPOPGLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDKL 138
Db 68 SVPIWERSPWVE-SEIGYTALSETSPRLFSHLPQLFPFKSFFSSKAKVIYLMRNPDRV 126

QY 139 VVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVQFWEHRMDSNVFLFKYED 198
Db 127 LVSGYFFWKNMKFIKKPKSWEEYFENFCQGTVVYGSWFDHGHGMPREEKNFLLLSYEE 186

QY 199 MHRDLVTWVEQLARFLGVSCDKAQLAEALTEHC---HQLVDQCCNAEALPVG----- 246
Db 187 LKQDTGRTIEKICQFLGKLTLEPELNLILKNSSFSQSMKENKMSYLLSDVYVDKAQLL 246

QY 247 -RGRVGLWKDIFTVSMNEKFDLVYKQKM 273
Db 247 RKGSGDWKNHFTVAQAEFDKLFQEKM 274

RESULT 70
US-10-851-921-12
; Sequence 12, Application US/10851921
; Publication No. US20040241737A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Khawaja, Xavier
; APPLICANT: Xu, Jun
; APPLICANT: Liang, JunJun
; TITLE OF INVENTION: METHODS FOR DIAGNOSING MOOD DISORDERS
; FILE REFERENCE: AM101227
; CURRENT APPLICATION NUMBER: US/10/851,921
; CURRENT FILING DATE: 2004-05-21
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-851-921-12

Query Match      23.2%; Score 355.5; DB 5; Length 285;
Best Local Similarity 31.0%; Pred. No. 8.2e-27;
Matches 83; Conservative 59; Mismatches 109; Indels 17; Gaps 5;

QY 22 FHGVRLLPP--FCRGKMEIAN-FVVRPSDVWIVTPKSGTSLQEVVLYVSGADPDEIG 78
Db 8 FEGIAFPTMGFRSETLRKVRDEFVIRDEVDVILTYPKSGTNWLABILCLMHKSGDAKWIQ 67

QY 79 LMNIDEQLPVLEYPOPGLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDKL 138
Db 68 SVPIWERSPWVE-SEIGYTALSETSPRLFSHLPQLFPFKSFFSSKAKVIYLMRNPDRV 126

QY 139 VVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVQFWEHRMDSNVFLFKYED 198
Db 127 LVSGYFFWKNMKFIKKPKSWEEYFENFCQGTVVYGSWFDHGHGMPREEKNFLLLSYEE 186

QY 199 MHRDLVTWVEQLARFLGVSCDKAQLAEALTEHC---HQLVDQCCNAEALPVG----- 246
```





; PRIOR APPLICATION NUMBER: 60/249,218  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,208  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,213  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,212  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,207  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,245  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,244  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,217  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,211  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,215  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,264  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,214  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,297  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/232,400  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/231,242  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/232,081  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/232,080  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,414  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,244  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/233,064  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/233,063  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/232,397  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/232,399  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/232,401  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/241,808  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,826  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,786  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,221  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/246,475  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/231,243  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/233,065  
; PRIOR FILING DATE: 2000-09-14

Query Match 20.3%; Score 310; DB 3; Length 105;  
Best Local Similarity 95.1%; Pred. No. 7.5e-23;  
Matches 58; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 165 FNDKLGVSWEHVFQEFWEHRMDSNVFLKDYEDMHRDLVTWVEQLARFLGVSCDKAOLE 224  
DB 45 FNDKLGVSWEHVFQEFWEHRMDSNVFLKDYEDMHRDLVTWVEQLARFLGVSCXIFOLE 104  
QY 225 A 225

Db 105 A 105  
RESULT 75  
US-09-795-926-2  
; Sequence 2, Application US/09795926  
; Patent No. US20020098486A1  
; GENERAL INFORMATION:  
; APPLICANT: Donoho, Gregory  
; APPLICANT: Hilbun, Erin  
; APPLICANT: Turner, C. Alexander Jr.  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Abuin, Alejandro  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Sands, Arthur T.  
; APPLICANT: Waite, D. Wade  
; APPLICANT: Wilganowski, Nathaniel L.  
; APPLICANT: Hu, Yi  
; APPLICANT: Kieke, James Alvin  
; APPLICANT: Potter, David George  
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND  
; FILE REFERENCE: POLYNUCLEOTIDES ENCODING THE SAME  
; CURRENT APPLICATION NUMBER: US/09/795,926  
; CURRENT FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 60/185,920  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 60/186,558  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: US 60/191,849  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 303  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-795-926-2  
Query Match 19.8%; Score 303.5; DB 3; Length 303;  
Best Local Similarity 26.7%; Pred. No. 1.5e-21;  
Matches 74; Conservative 61; Mismatches 99; Indels 43; Gaps 9;  
QY 19 YFEHGVRLP-PFCRGK-MEEIANPPVRPSPVWIVTPKSGTS-----LLOEVVVLVSQGA 72  
DB 27 FTYOGIPYPTMTCTSETFQALDTFEARHDDIVLASYPKCSNWILHIVSELIYAVSK-- 84  
QY 73 DPDEIGLMNIDEQLFPVLEYPQ-PGLDI-----IKELTSPLIKSHLPYRFLPSDLHN 123  
DB 85 -----KKYKYPFPVLECGDSEKQYQRMKGFPSPRILATHLHYDKLPGSIFE 130  
QY 124 GDSKVIYMARNPDLVSYQFHRSLRTWMSYRGTFQFCRFRPMNDKLGYSWFEHVQSEFW 183  
DB 131 NKAKILVIFRNPKDTAVSFLHFNDDVPDIPSYGSMDEFFRQFMKGQVSGWGRYDFFAINNN 190  
QY 184 EHRMDSNVFLKDYEDMHRDLVTWVEQLARFLGVSCDKAOLEALTEHCHQLVDQCCNABAL 243  
DB 191 KHLGDGVKFIYEDLKENLAAGIKQIAEFLGFFUTGTGQIQTISV---QSTFQAMRAKSQ 247  
QY 244 ----PVG-----RGRVGLWKDIFTVSMNEKFDLVYKQ 271  
DB 248 DTHGAVGPFPLFRKGEVGDWKNLFSIQNQEMDEKPE 284  
RESULT 76  
US-10-364-774-2  
; Sequence 2, Application US/10364774  
; Publication No. US20030144497A1  
; GENERAL INFORMATION:  
; APPLICANT: Donoho, Gregory  
; APPLICANT: Hilbun, Erin  
; APPLICANT: Turner, C. Alexander Jr.  
; APPLICANT: Friedrich, Glenn

```
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: POLYNUCLEOTIDES ENCODING THE SAME
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/10/364,774
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 303
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774-2

Query Match      19.8%; Score 303.5; DB 4; Length 303;
Best Local Similarity 26.7%; Pred. No. 1.5e-21;
Matches 74; Conservative 61; Mismatches 99; Indels 43; Gaps 9;

Qy 19 YFEFHGVRLP-PFCRGK-MEEIANFPVRPSPDVWIVTPKSGTS-----LLQEVVYLVSOGA 72
Db 27 FFTYQGIPIPTMCTSETFFQALDTEARHDDIVLASYPKCGSNWILHIVSELIYAVSK-- 84

Qy 73 DPDEIGLNMNIDEQLPVLEYPO-PGLDI-----IKELTSPLRIKSHLPYRFLPSDLHN 123
Db 85 -----KKYKYPFPVLECGDSEKYOQRMKGFPSPRILATHLHYDKLPGSIFE 130

Qy 124 GDSKVIYARNPKDLVSVYQFHRSLRTMSVGRGTQBFQECRRFMDKLGYSWFEHVQBFW 183
Db 131 NKAKILVIFRNPKOTAVSFLHFHNDVPDIPSYGSWDEFFRQPMKGQVSWGRYFDFAINWN 190

Qy 184 EHRMDSNVFLKYEDMHRDLVTWVQLARFLGVSCDKAQLAALTEHCHQLVDQCCNAEAL 243
Db 191 KHLGDNVKFLYLEDLKENLAAGIKQIAEFLGFLTGQIQTISV---QSTFQAMRAKSQ 247

Qy 244 ----PVG-----RGRVGLWKDIFTVSMNEKFDLVYKQ 271
Db 248 DTHGAVGPFPRKGEVGDWKNLFSIEIQNQEEMDEKFE 284

RESULT 77
US-10-072-012-664
; Sequence 664, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zethusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
```

```
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 664
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-664

Query Match      19.8%; Score 303.5; DB 4; Length 303;
Best Local Similarity 26.7%; Pred. No. 1.5e-21;
Matches 74; Conservative 61; Mismatches 99; Indels 43; Gaps 9;

Qy 19 YFEFHGVRLP-PFCRGK-MEEIANFPVRPSPDVWIVTPKSGTS-----LLQEVVYLVSOGA 72
Db 27 FFTYQGIPIPTMCTSETFFQALDTEARHDDIVLASYPKCGSNWILHIVSELIYAVSK-- 84

Qy 73 DPDEIGLNMNIDEQLPVLEYPO-PGLDI-----IKELTSPLRIKSHLPYRFLPSDLHN 123
Db 85 -----KKYKYPFPVLECGDSEKYOQRMKGFPSPRILATHLHYDKLPGSIFE 130

Qy 124 GDSKVIYARNPKDLVSVYQFHRSLRTMSVGRGTQBFQECRRFMDKLGYSWFEHVQBFW 183
Db 131 NKAKILVIFRNPKOTAVSFLHFHNDVPDIPSYGSWDEFFRQPMKGQVSWGRYFDFAINWN 190

Qy 184 EHRMDSNVFLKYEDMHRDLVTWVQLARFLGVSCDKAQLAALTEHCHQLVDQCCNAEAL 243
Db 191 KHLGDNVKFLYLEDLKENLAAGIKQIAEFLGFLTGQIQTISV---QSTFQAMRAKSQ 247

Qy 244 ----PVG-----RGRVGLWKDIFTVSMNEKFDLVYKQ 271
Db 248 DTHGAVGPFPRKGEVGDWKNLFSIEIQNQEEMDEKFE 284

RESULT 78
US-10-468-125-3
; Sequence 3, Application US/10468125
; Publication No. US20040082061A1
; GENERAL INFORMATION:
; APPLICANT: ASTROMOFF, Anna
```

```

; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: DING, Li
; APPLICANT: DUGGAN, Brendan M.
; APPLICANT: FORTSYTHE, Ian J.
; APPLICANT: GIETZEN, Kimberly J.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: LU, Yan
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: RING, Huijun Z.
; APPLICANT: SANJANWALA, Madhusudan
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: WARREN, Bridget A.
; APPLICANT: XU, Yuming
; APPLICANT: YUE, Henry
; APPLICANT: ZSEBARJADIAN, Yeganeh
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0363 USN
; CURRENT APPLICATION NUMBER: US/10/468,125
; CURRENT FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: PCT/US02/04918
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,643
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/271,332
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/276,767
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: US 60/282,077
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/285,447
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/287,060
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/288,543
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7491172CD1
US-10-468-125-3

Query Match 19.8%; Score 303.5; DB 4; Length 303;
Best local Similarity 26.7%; Pred. No. 1.5e-21;
Matches 74; Conservative 61; Mismatches 99; Indels 43; Gaps 9

Qy 19 YFEPHGVRLP-PFCRGK-MESIANFPVRPSPDWIVTYPKSGTS-----LQEVVYLVSGA 72
Db 27 FFTVQGIPIYPTMTSETTFQALDTFEARHDDIVLASYPKCSNWILHIVSELIVAVSK-- 84

Qy 73 DPDEIGLNMNIDEQLPLVEYPO-PGLDI-----IKELTGPRLIKSHLPVRFPSDLHN 123
Db 85 -----KKKYKPPFVLECGDSEKYSQRMKGFPPSRILATHLHYDKLPGSIFE 130

Qy 124 GDSKVIYWARNPDKLVISYQFHSRLRPMVSRGTFQECRRFMNDKLGYSWFHVFQSF 183
Db 131 NKAKILVIFRNPDKTAVSFLRHFHNDVPDIPSYGSWDEFFRPMKQGVSWGRYFDFAINW 190

Qy 184 EHRMDSNVLFUKYEDMRHDLVTWVEQLARFLGVSCKQLAEALTEHCHQLVDQCCNABAL 243
Db 191 KHLGDGNVKFLYEDLKENLAAGIKQIAEFLGFFLTGSIQTISV---QSTFQAMRAKSO 247

Qy 244 -----PVG-----RGRVGLWKDIFTVSMNEKFDLVYKQ 271
Db 248 DTHGAVGPFPLPRKGEVGDWKNLFSFIQNMDEKFKF 284

```







```
; CURRENT APPLICATION NUMBER: US/11/134,241
; CURRENT FILING DATE: 2005-05-20
; PRIOR APPLICATION NUMBER: US/10/364,774
; PRIOR FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 265
; TYPE: PRT
; ORGANISM: homo sapiens
US-11-134-241-6

Query Match      19.1%; Score 292.5; DB 6; Length 265;
Best Local Similarity 27.1%; Pred. No. 1.6e-20;
Matches 70; Conservative 57; Mismatches 90; Indels 41; Gaps 7;

QY 36 EBIANPVRPSDVWVITPKSGTS-----LLQEVVLVSGADPDDEIGLNMNIDEQLPVLEY 91
Db 8 QALDTFEARHDDIVLASYPKCGSNMILHIVSELIVAVSK-----KKYKY 51

QY 92 PQ-PGLDI-----IKELTSPRIKSHLPYRFLPSPDLHNGDSKVIYMARNPKDLVVS 142
Db 52 PEPFVLECGDSKYQRMKGFPSPRIATHLHVDKLPGSIFENKAKILVIFRNPKDTAVSF 111

QY 143 YQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVEHVEFWEHRMDSNVFLFLKYEDMHRD 202
Db 112 LHFHNDVPDIPSGVSGWDEFFRFQFMKGQVSGRYDFDFAINWNKHLDGDNVVKFLYEDLKEN 171

QY 203 LVTWVEQLARFLGVSCDKAQLEALTEHCHQVLDVQCCNAEAL-----PVG-----RGRVGLW 253
Db 172 LAAGIKQIAEFLGFLTGEQIQTISV---QSTFQAMRAKSDQTHGAVGPFPLFRKGEVGDW 228

QY 254 KDIFTVMNEKFDLVYKQ 271
Db 229 KNLFSEIQNQEWDKFEKE 246

RESULT 85
US-10-072-012-665
; Sequence 665, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Futak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 665
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Gallus gallus
US-10-072-012-665

Query Match      18.9%; Score 289.5; DB 4; Length 312;
Best Local Similarity 25.4%; Pred. No. 3.9e-20;
Matches 69; Conservative 68; Mismatches 122; Indels 13; Gaps 6;

QY 20 PEFHGVRLP-PFCRGKM-EETANFPVRPSDVWVITPKSGTSLLOEVVLVSGADPDEI 77
Db 29 PSYKGVLPVALCSPEVFRAMESFEARSDDDVILAGYPKSGTNWVGQILSDLVATFEKERL 88

QY 78 GLMNI-DEQLPVLEYPOPG---LDIIKELTSPRIKSHLPYRFLPSPDLHNGDSKVIYMA 132
Db 89 BEKSVNDEELEEFPYLEIGDTEKYERMKKLPSRRVILTHLSPEKLPKIFNNAKILLI 148

QY 133 RNPDLVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVEHVEFWEHRMDSNV 192
Db 149 RNPDIATSPFHSNRWSALPSYETWDDFFTAFTKMPWGSYFNLYSEWNKYAADENV 208

QY 193 FLKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHC--HQLVDQCCNAE----ALPVG 246
Db 209 TITVEELKENQTLGVKNIASPFGISLTGEELRSVIERSFSQSMKENSILKTHGALGMLFR 268

QY 247 RGRVGLWKDIFTVMNEKFDLVYKQKMGKCDL 278
Db 269 KGVGSDWKNLFPNEQNEKMDKVFERTARTKL 300

RESULT 86
US-10-072-012-250
; Sequence 250, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
```

```
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 250
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-072-012-250

Query Match      18.4%; Score 281.5; DB 4; Length 303;
Best Local Similarity 26.0%; Pred. No. 2.4e-19;
Matches 72; Conservative 61; Mismatches 101; Indels 43; Gaps 9;

Qy 19 YPEFHGVRLP-PFCRGK-MEEIANTFVPSPDVWIVTPKSGTS-----LLQEVVYLVSGA 72
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
27 FPTYQIGYPITMCTSETFQALDTFEARHDDIVLASYPKCGSNWILHIVSELIVAVSK-- 84
Qy 73 DPDEIGLANNIDQLVLEVPQ-PGLDI-----IKELTSPLRIKSHLPYRFLPSDLHN 123
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
85 -----KKYKYPEPVLECGDSEKQKMGFPSPRIATHLHYDKLPGSIFE 130
Qy 124 GDSKVIYMARNPKOLVSYQYQPHRSRLTMSYRGTFQFCRRFPMNDKLGYSWFHVBQFW 183
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
131 NKAKILVIFRNPKDTAVSFLPHNDVPDIPSYGSDWDFPRQPMVFLVSWGRYDFDAINWN 190
Qy 184 EHRMDSNVLFUKYEDMHRDLVTMWQLARFLGVSCDKAQLAALTEHCHQLVDVQCCNAAL 243
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
191 KHLGDGNVKFILDYELKELVRLGKQIAFLGFLFTGTGEIQITISV---QSTFQAMRAKSQ 247
Qy 244 ----PVG-----RGRVGLWKDIFTVSMNEKFDLVYKQ 271
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
248 DTHGAVGPFPLFRKGVADWKNLFSFIQIQWEMDEKFE 284
```

RESULT 87  
US-11-096-568A-1016

```
; Sequence 1016, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 1016
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(345)
; OTHER INFORMATION: Ceres Seq. ID no. 15218612
US-11-096-568A-1016
```

```
Query Match      18.1%; Score 276.5; DB 6; Length 345;
Best Local Similarity 27.6%; Pred. No. 9.1e-19;
Matches 89; Conservative 44; Mismatches 123; Indels 67; Gaps 12;

Qy 10 STPGEFESK--YFEPFHGV-----RLLPPFCR-----GKMBEIANFP 42
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
12 STPEADEAKIYDEARRVVSTYETVPSPGTLQDYCRHPSCWCITLPTWSSMVAEQHFE 71
Qy 43 VRPSDVWIVTPKSGTSLLOEVVYLVSGADPDPIGLM-----NIDQLPVLE---YPQ 93
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
72 ARGTDVLLVTPKSGTTWIKALLVAAAHRTDTSLSILQLASHNSHQVLPFLAQVYTK 131
Qy 94 PGLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVSYQYQPHRSRLTMS 153
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
132 DQIPDLSLPAPRLFATHIPAESLPSPSVVAGSKVYVLCRDPKDCFCVSLMHPMKNKFTPD 191
Qy 154 YRGTFQFCRRFMDKLG---YGSWFHGVQFWEHRMD--SNVLFLKYEDMHRDLVTWVE 208
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
192 IDEAHGRFCE-----GVSLYGFWEHVLVSYWVHVDPRQGVFLTYEELSADPLGQLR 244
Qy 209 QLARFLGVSCDKAQLA-----LTEHC-----HQLVDVQCCNAE-----ALPVG---RGR 249
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
245 RLAEFIGRPPTPGEGAGVDREIAEACAMKSMVQVNSRRTTEIVEMPINGIFFRRGV 304
Qy 250 VGLWKDIFTVSMNEKFDLVYKQ 272
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
305 VGDWNTNLTPEMAGRIDEITKSK 327
```

RESULT 88  
US-09-854-122-20  
; Sequence 20, Application US/09854122  
; Patent No. US20020016980A1  
; GENERAL INFORMATION:  
; APPLICANT: SMITH, ROBERT  
; APPLICANT: ALBERTE, RANDALL S.  
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA  
; FILE REFERENCE: PHA-007.01  
; CURRENT APPLICATION NUMBER: US/09/854,122  
; CURRENT FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: 60/202,529  
; PRIOR FILING DATE: 2000-05-10  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 320  
; TYPE: PRT  
; ORGANISM: Flaveria bidentis  
US-09-854-122-20

Query Match 18.0%; Score 275.5; DB 3; Length 320;  
Best Local Similarity 27.5%; Pred. No. 1e-18;  
Matches 80; Conservative 63; Mismatches 111; Indels 37; Gaps 11;



US-11-096-568A-1018

Query Match 17.9%; Score 274; DB 6; Length 282;  
Best Local Similarity 28.9%; Pred. No. 1.2e-18;  
Matches 77; Conservative 42; Mismatches 107; Indels 40; Gaps 9;  
  
QY 40 NFPVRPSDVWIVTPKSGTSLQEVVYLVSGQADPDEIGLM-----NIDQLPVLE--- 90  
DB 6 HFEARGTDVLLVTPKSGTTWIKALLYAAARTDDTSSILRLQASHNSHQVFPLEAQV 65  
QY 91 YPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVSVYQFHRSLR 150  
DB 66 YTKQIPDLSSLAPRLPETHIPAESLPPSVVASCCKVYLCRDPKDCVSLWHFMNKF 125  
QY 151 TMSYRGTFQECRRPMNDKLG---YGSWFHVFQEFWEHRMD--SNVFLKYEDMHRDLVT 205  
DB 126 PWDIDEAHGRFCE-----GVSLYGPFWEHVLVSYRWHVDRPGQVFLTYEELSADPLG 178  
QY 206 MVEQLARFLGVSCDKAQLA-----LTHC-----HOLVQCCNAE---ALPVG--- 246  
DB 179 QLRLAEFIGRPFTGEGEAGVDREIAEACAMKSMVNOEVNQSRTEIVEMPINGIFFR 238  
QY 247 RGRVGLWKDIFTVSMNEKFDLVYKQK 272  
DB 239 RGVGDWNTNLTPEMAGRIDETITKSK 264

RESULT 92

US-11-096-568A-1017  
; Sequence 1017, Application US/11096568A  
; Publication No. US20060048240A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nickolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096,568A  
; CURRENT FILING DATE: 2005-04-01  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 1017  
; LENGTH: 286  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(286)  
; OTHER INFORMATION: Ceres Seq. ID no. 15218613  
US-11-096-568A-1017

Query Match 17.9%; Score 274; DB 6; Length 286;  
Best Local Similarity 28.9%; Pred. No. 1.2e-18;  
Matches 77; Conservative 42; Mismatches 107; Indels 40; Gaps 9;  
  
QY 40 NFPVRPSDVWIVTPKSGTSLQEVVYLVSGQADPDEIGLM-----NIDQLPVLE--- 90  
DB 10 HFEARGTDVLLVTPKSGTTWIKALLYAAARTDDTSSILRLQASHNSHQVFPLEAQV 69  
QY 91 YPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVSVYQFHRSLR 150  
DB 70 YTKQIPDLSSLAPRLPETHIPAESLPPSVVASCCKVYLCRDPKDCVSLWHFMNKF 129  
QY 151 TMSYRGTFQECRRPMNDKLG---YGSWFHVFQEFWEHRMD--SNVFLKYEDMHRDLVT 205  
DB 130 PWDIDEAHGRFCE-----GVSLYGPFWEHVLVSYRWHVDRPGQVFLTYEELSADPLG 182  
QY 206 MVEQLARFLGVSCDKAQLA-----LTHC-----HOLVQCCNAE---ALPVG--- 246  
DB 183 QLRLAEFIGRPFTGEGEAGVDREIAEACAMKSMVNOEVNQSRTEIVEMPINGIFFR 242  
QY 247 RGRVGLWKDIFTVSMNEKFDLVYKQK 272  
DB 243 RGVGDWNTNLTPEMAGRIDETITKSK 268

RESULT 93

US-10-437-963-198983  
; Sequence 198983, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)78  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 198983  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_94591C.1.pep  
US-10-437-963-198983

Query Match 17.6%; Score 269.5; DB 4; Length 336;  
Best Local Similarity 27.3%; Pred. No. 4.4e-18;  
Matches 73; Conservative 54; Mismatches 97; Indels 43; Gaps 9;

QY 40 NFPVRPSDVWIVTPKSGTSLQEVVYLVSGQADPDEIGLMNIDEQLPVLEYP----- 92  
DB 46 SFQRPSPDILLSSFPKCGTTLWKALAFATLNRTSTPPS-----DEHHPLLEHNPDLVG 99  
QY 93 ----QPLDIIKELTSPRLIKSHLPYRFLPSDLHNGDS--KVIYMARPKDLVSVYQFH 146  
DB 100 FLEIYPKLELYESLPSPRLISTHLPSMLPHRTRQETGCLVYIYRDPKAMVS--MWH 157  
QY 147 RSLTMSYRGTFQECRRPMNDKLGYSWFHVFQEFWEHRM--DSNVFLKYEDMHRDLV 204  
DB 158 QNKKEKNRLTTFEEMDFMFCGRCVGPQWCHAGEYWDQSQAPEKVLFLMYEDLLQDTV 217  
QY 205 TMVEQLARFLGVSCDKAQLAELTEHCHQLVQDCC--NAAALPYG----- 246  
DB 218 GNLTLEAFMG--CGFSRQBEDDGIQQIVELCSLNLKLNLVNKS GTTLLIGISKDGFFR 275  
QY 247 RGRVGLWKDIFTVSMNEKFDLVYKQK 273  
DB 276 KGGTGDWSNHSMPMAARLDKIVKERL 302

RESULT 94

US-11-096-568A-7712  
; Sequence 7712, Application US/11096568A  
; Publication No. US20060048240A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nickolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096,568A  
; CURRENT FILING DATE: 2005-04-01  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 7712  
; LENGTH: 329  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(329)  
; OTHER INFORMATION: Ceres Seq. ID no. 15174866

US-11-096-568A-7712

Query Match 17.5%; Score 268.5; DB 6; Length 329;  
Best Local Similarity 28.1%; Pred. No. 5.4e-18;  
Matches 78; Conservative 45; Mismatches 106; Indels 49; Gaps 9;  
  
QY 41 FVPRSDVMIVTPKSGTSLQEVVVLVSQ----GADPDEIGLMNIDEOLPVLEYP--QP 94  
Db 63 FVRADDDIILATQPKCGITLWKALAFITANRSRYGSDHPLLTRHPQHVPVPIEIPGAGP 122  
  
QY 95 GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDKLVVSYQFHRSLRTMSY 154  
Db 123 NHTDHTLPPRLATHMPMSLLPPGTRSVGCGEVVYLCRDPKDALVSRHLFENK----AF 178  
  
QY 155 RGT-----FQEFRRFMNDKLGYSWFEHVOEFWEHRM--DSNVLFKLYEDMHRDLVT 205  
Db 179 QGTNLSDMNAFMSFCGFS----PYGPFWDHCLGYWRESIARPDNVLFKLYEIKSDPVP 234  
  
QY 206 MVEQLARFLGVSCDKAQLEALTEHCHQVLDQCCNAEAL-----PVG----- 246  
Db 235 VVRKLAKFLGVPLTE---EESSGVAQEVVVKLCSPETLTSLQVNVQGVVGRHGDKIFMSNS 291  
  
QY 247 ----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
Db 292 VFYRKEGVDWANHSHVWGEEELDRIVQOKLEGSLGVF 329

RESULT 95

US-11-096-568A-7711  
; Sequence 7711, Application US/11096568A  
; Publication No. US20060048240A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nikolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096,568A  
; CURRENT FILING DATE: 2005-04-01  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 7711  
; LENGTH: 330  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)..(330)  
; OTHER INFORMATION: Ceres Seq. ID no. 15174865  
US-11-096-568A-7711

Query Match 17.5%; Score 268.5; DB 6; Length 330;  
Best Local Similarity 28.1%; Pred. No. 5.4e-18;  
Matches 78; Conservative 45; Mismatches 106; Indels 49; Gaps 9;  
  
QY 41 FVPRSDVMIVTPKSGTSLQEVVVLVSQ----GADPDEIGLMNIDEOLPVLEYP--QP 94  
Db 64 FVRADDDIILATQPKCGITLWKALAFITANRSRYGSDHPLLTRHPQHVPVPIEIPGAGP 123  
  
QY 95 GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDKLVVSYQFHRSLRTMSY 154  
Db 124 NHTDHTLPPRLATHMPMSLLPPGTRSVGCGEVVYLCRDPKDALVSRHLFENK----AF 179  
  
QY 155 RGT-----FQEFRRFMNDKLGYSWFEHVOEFWEHRM--DSNVLFKLYEDMHRDLVT 205  
Db 180 QGTNLSDMNAFMSFCGFS----PYGPFWDHCLGYWRESIARPDNVLFKLYEIKSDPVP 235  
  
QY 206 MVEQLARFLGVSCDKAQLEALTEHCHQVLDQCCNAEAL-----PVG----- 246  
Db 236 VVRKLAKFLGVPLTE---EESSGVAQEVVVKLCSPETLTSLQVNVQGVVGRHGDKIFMSNS 292  
  
QY 247 ----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
Db 293 VFYRKEGVDWANHSHVWGEEELDRIVQOKLEGSLGVF 330

RESULT 96

US-11-096-568A-7710  
; Sequence 7710, Application US/11096568A  
; Publication No. US20060048240A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nikolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096,568A  
; CURRENT FILING DATE: 2005-04-01  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 7710  
; LENGTH: 340  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)..(340)  
; OTHER INFORMATION: Ceres Seq. ID no. 15174864  
US-11-096-568A-7710

Query Match 17.5%; Score 268.5; DB 6; Length 340;  
Best Local Similarity 28.1%; Pred. No. 5.7e-18;  
Matches 78; Conservative 45; Mismatches 106; Indels 49; Gaps 9;  
  
QY 41 FVPRSDVMIVTPKSGTSLQEVVVLVSQ----GADPDEIGLMNIDEOLPVLEYP--QP 94  
Db 74 FVRADDDIILATQPKCGITLWKALAFITANRSRYGSDHPLLTRHPQHVPVPIEIPGAGP 133  
  
QY 95 GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDKLVVSYQFHRSLRTMSY 154  
Db 134 NHTDHTLPPRLATHMPMSLLPPGTRSVGCGEVVYLCRDPKDALVSRHLFENK----AF 189  
  
QY 155 RGT-----FQEFRRFMNDKLGYSWFEHVOEFWEHRM--DSNVLFKLYEDMHRDLVT 205  
Db 190 QGTNLSDMNAFMSFCGFS----PYGPFWDHCLGYWRESIARPDNVLFKLYEIKSDPVP 245  
  
QY 206 MVEQLARFLGVSCDKAQLEALTEHCHQVLDQCCNAEAL-----PVG----- 246  
Db 246 VVRKLAKFLGVPLTE---EESSGVAQEVVVKLCSPETLTSLQVNVQGVVGRHGDKIFMSNS 302  
  
QY 247 ----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
Db 303 VFYRKEGVDWANHSHVWGEEELDRIVQOKLEGSLGVF 340

RESULT 97

US-10-767-701-42402  
; Sequence 42402, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 42402  
; LENGTH: 344  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C44289\_1.pep  
US-10-767-701-42402

Query Match 17.5%; Score 267.5; DB 4; Length 344;  
Best Local Similarity 28.4%; Pred. No. 7.2e-18;  
Matches 78; Conservative 47; Mismatches 109; Indels 41; Gaps 10;







**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2006, 10:59:32 ; Search time 300 Seconds  
(without alignments)  
875.681 Million cell updates/sec

Title: US-10-768-158-2

Perfect score: 1930

Sequence: 1 MAESEATPSTPGFESKYF.....FDLVYKQMGKCDLTFDFYL 284

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt 7.2.1\*

1: uniprot\_sprot:\*

2: uniprot\_trenbl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1530	100.0	284	1	ST4A1_HUMAN
2	1525	99.7	284	2	QSR4X3_PONPY
3	1508	98.6	284	2	Q865L7_RABIT
4	1506	98.4	284	1	ST4A1_MOUSE
5	1506	98.4	284	1	ST4A1_RAT
6	1506	98.4	284	2	Q3TXI5_MOUSE
7	1483	96.9	284	2	QBQV8_MOUSE
8	1385	90.5	284	2	Q68EV3_XENLA
9	1382	90.3	284	2	Q52KW3_XENLA
10	1313	85.8	284	2	Q4SRF8_TETNG
11	995	65.0	218	2	Q5RCAB_PONPY
12	495	32.4	298	2	Q9RIS_MOUSE
13	489	32.0	291	1	ST1A1_MOUSE
14	480	31.4	296	1	ST1C1_MOUSE
15	476	31.1	296	2	Q3UNA9_MOUSE
16	472	30.8	295	1	ST1A3_HUMAN
17	467	30.5	296	1	STK1_RAT
18	467	30.5	296	2	Q3ZAV3_RAT
19	466	30.5	296	2	Q64ZG8_RAT
20	465	30.4	296	2	Q569D0_RAT
21	464	30.3	295	2	Q2TAB3_HUMAN
22	463	30.3	298	2	Q58CV8_BOVIN
23	459.5	30.0	296	1	ST1B1_CHICK
24	458	29.9	291	1	ST1A1_RAT
25	458	29.9	291	2	Q548D2_RAT
26	456	29.8	295	1	ST1A2_HUMAN
27	456	29.8	295	2	Q6Y0X5_PIG
28	456	29.8	296	1	STK2_RAT
29	453.5	29.6	296	2	Q95JC6_TRIVU
30	451	29.5	295	1	ST1A1_HUMAN
31	451	29.5	295	2	Q2NL71_HUMAN

32	451	29.5	295	2	Q3ZC30_BOVIN
33	450	29.4	295	1	ST1A1_CANFA
34	449	29.3	295	2	Q95KM3_RABIT
35	449	29.3	296	1	ST1C1_HUMAN
36	449	29.3	302	1	ST1C2_HUMAN
37	447	29.2	295	2	Q95MP8_PIG
38	446.5	29.2	299	1	ST1B1_RAT
39	446	29.2	295	1	ST1A1_MACFA
40	446	29.2	295	1	ST1E1_BOVIN
41	443.5	29.0	296	1	ST1B1_HUMAN
42	443	29.0	296	1	ST1B1_CANFA
43	442.5	28.9	296	1	ST1C1_RABIT
44	442	28.9	292	2	Q9XT98_RABIT
45	440	28.8	297	2	Q8GP49_XENLA
46	436.5	28.5	296	2	Q3T0Y3_BOVIN
47	435	28.4	295	1	ST1E2_RAT
48	435	28.4	295	1	ST1E6_RAT
49	434.5	28.4	299	1	ST1S1_BRARE
50	434	28.4	294	1	ST1E1_HUMAN
51	434	28.4	294	2	Q53X91_HUMAN
52	432.5	28.3	301	2	Q6DHG7_BRARE
53	432	28.2	295	2	Q9D566_MOUSE
54	431.5	28.2	287	2	Q5HZV7_XENTR
55	431	28.2	295	1	ST1E1_MOUSE
56	431	28.2	295	2	Q8JZX7_MOUSE
57	430	28.1	295	2	Q99ND5_RAT
58	429.5	28.1	299	1	ST1B1_MOUSE
59	429	28.0	263	2	Q91W19_MOUSE
60	429	28.0	295	1	ST1E3_RAT
61	429	28.0	295	2	Q3T0S9_BOVIN
62	428	28.0	307	2	Q90WR6_CHICK
63	427.5	27.9	301	1	ST1S3_BRARE
64	427.5	27.9	301	1	ST2S2_BRARE
65	427.5	27.9	302	2	Q4FZP1_XENLA
66	423	27.6	294	2	Q2XV97_MACFA
67	422	27.6	304	2	Q61M16_HUMAN
68	421.5	27.5	304	2	Q8XZC1_BRARE
69	421	27.5	295	1	ST1E1_RAT
70	419.5	27.4	285	2	Q3HYK0_STRPU
71	419.5	27.4	304	1	SUAC_RAT
72	419.5	27.4	304	2	Q5M8B5_RAT
73	416.5	27.2	287	2	Q6PG4_XENTR
74	414.5	27.1	276	2	Q88EV4_XENLA
75	414	27.1	295	2	Q95JCS_ORNAN
76	413	27.0	295	2	Q35401_MOUSE
77	413	27.0	303	2	Q5EAW0_XENLA
78	412.5	27.0	304	2	Q80VR3_MOUSE
79	412	26.9	294	1	ST1A1_BOVIN
80	412	26.9	294	2	Q7S293_XENLA
81	412	26.9	295	2	Q95JD6_CANFA
82	412	26.9	295	2	Q3UZ26_MOUSE
83	412	26.9	295	2	Q6NZD1_MOUSE
84	412	26.9	308	2	Q921G0_RAT
85	412	26.9	309	2	Q9R2C2_MOUSE
86	411	26.8	305	2	Q66KW4_XENLA
87	410.5	26.8	297	2	Q5HZU0_XENTR
88	408.5	26.7	304	1	ST1E1_CAPOU
89	408.5	26.7	306	2	Q70262_MOUSE
90	405.5	26.5	309	2	Q4RG19_TETNG
91	405	26.5	316	2	Q9VHH0_DROME
92	404.5	26.4	287	2	Q8DJL6_XENLA
93	402	26.3	293	2	Q5M7C2_XENLA
94	401.5	26.2	293	2	Q49IK6_BRARE
95	401	26.2	304	2	Q61MI5_HUMAN
96	395.5	25.8	307	2	Q53SG4_HUMAN
97	393.5	25.7	301	1	ST1A1_RABIT
98	393.5	25.7	304	2	Q6DJ68_XENTR
99	393.5	25.7	304	2	Q7ZYH0_XENLA
100	392.5	25.7	299	2	Q6GL93_XENTR

ALIGNMENTS

RESULT 1  
ST4A1 HUMAN  
ID ST4A1\_HUMAN STANDARD; PRT; 284 AA.  
AC Q9BR01; O43728;  
DC 02-AUG-2002, integrated into UniProtKB/Swiss-Prot.  
DT 02-AUG-2002, sequence version 2.  
DT 07-FEB-2006, entry version 33.  
DE Sulfotransferase 4A1 (EC 2.8.2.-) (Brain sulfotransferase-like  
DE protein) (HBR-STL) (HBR-STL-1) (Nervous system sulfotransferase  
DE (NST)).  
GN Name=SULT4A1; Synonyms=SULTX3;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI TaxID=9606;  
[1]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 1), AND CHARACTERIZATION.  
RC TISSUE=Brain;  
RX MEDLINE=20164074; PubMed=10698717; DOI=10.1042/0264-6021.3460857;  
RA Falany C.N., Xie X., Wang J., Ferrer J., Falany J.L.;  
RT "Molecular cloning and expression of novel sulphotransferase-like  
RT cDNAs from human and rat brain.";  
RL Biochem. J. 346:857-864(2000).  
[2]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 1), AND CHARACTERIZATION.  
RC TISSUE=Brain;  
RX MEDLINE=22035352; PubMed=12039030; DOI=10.1016/S0378-1119(02)00431-6;  
RA Sakakibara Y., Suiko M., Pai T.G., Nakayama T., Takami Y.,  
RA Katsufuchi J., Liu M.-C.;  
RT "Highly conserved mouse and human brain sulfotransferases: molecular  
RT cloning, expression, and functional characterization.";  
RL Gene 285:39-47(2002).  
[3]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).  
RC TISSUE=Brain;  
RA Martin S.C., Farb D.H.;  
RT "Molecular identification of a human nervous system cytoplasmic  
RT sulfotransferase, NST.";  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
[4]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).  
RC TISSUE=Brain;  
RA Walther S.E., Raftogiannis R.B.;  
RT "Molecular and physical characterization of human SULT4A1,  
RT representing a novel cytosolic sulfotransferase family.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[5]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 2).  
RC TISSUE=Testis;  
RA Goward M.E., Huckle E.J.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
[6]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RX PubMed=15461802; DOI=10.1106/gb-2004-5-10-r84;  
RA Collins J.E., Wright C.L., Edwards C.A., Davis M.P., Grinham J.A.,  
RA Cole C.G., Goward M.E., Aguado B., Mallia M., Mokrab Y., Huckle E.J.,  
RA Beare D.M., Dunham I.;  
RT "A genome annotation-driven approach to cloning the human ORFeome.";  
RL Genome Biol. 5:RESEARCH84.1-RESEARCH84.11(2004).  
[7]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RX MEDLINE=20057165; PubMed=10591208; DOI=10.1038/990031;  
RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,  
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,  
RA Bagnoley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,  
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,  
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,  
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,  
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,  
RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,  
RA Evans K.L., Fay J.M., Fleming K., French L., Garner A.A.,

Gilbert J.G.R., Goward M.E., Grahame D.V., Griffiths M.N.D., Hall C.,  
Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,  
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,  
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,  
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.T.,  
RA Mcclay J., McLaren S., McMurray A.M., Milne S.A., Mortimore B.J.C.T.,  
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,  
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,  
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,  
RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,  
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,  
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,  
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,  
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,  
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,  
RA Roe B.A., Chen F., Chu L., Crabtree J., Aoki N., Lao H.I.,  
RA Doxman A., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,  
RA Lewis J., Lewis S., Lin S.-P., Loh P., Mala J.E., Nguyen T., Pan H.,  
RA Phan S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,  
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,  
RA Zhan M., Zhang G., Chisoso S., Murray J., Miller N., Minx P.,  
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,  
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,  
RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,  
RA Scheet P., Walker C., Wamaley A., Wohldmann P., Pepin K., Nelson J.,  
RA Korf I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,  
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,  
RA Budarf M.L., Mcdermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,  
RA Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dumanaki J.P.,  
RA Peyrard M., Kedra D., Seroussi E., Francon I., Tapia I., Bruder C.E.,  
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,  
RA Khan A.S., Lane L., Tilahun Y., Wright H.;  
RT "The DNA sequence of human chromosome 22.";  
RL Nature 402:489-495(1999).  
[8]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).  
RC TISSUE=Fetal brain, Hippocampus, and Hypothalamus;  
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Berge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haiech F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).  
CC -!- FUNCTION: May catalyze the sulfate conjugation of many drugs,  
CC xenobiotic compounds, hormones, and neurotransmitters. Displays  
CC activity towards L-triiodothyronine, thyroxine, estrone, p-  
CC nitrophenol, 2-naphthylamine, and 2-naphthol. May have a role in the  
CC sulfation of drugs and neurotransmitters in the CNS.  
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q9BR01-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q9BR01-2; Sequence=VSP\_006304;  
CC Note=No experimental confirmation available;  
CC -!- TISSUE SPECIFICITY: Highly expressed in the cerebral cortex and  
CC frontal lobe, slightly less in the cerebellum, occipital and

CC temporal lobes, relatively low in the medulla and putamen, and  
CC lowest in the spinal cord. No expression detected in the pancreas  
CC (Ref.1). Highly expressed in fetal brain and occipital lobe,  
CC slightly less in the whole brain, frontal lobe, hippocampus, and  
CC lung, very low expression in cerebellum, medulla oblongata,  
CC temporal lobe, testis, kidney and appendix (Ref.2).  
CC -!- SIMILARITY: Belongs to the sulfotransferase family.  
CC  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC  
CC  
CC EMBL: AF188698; AAF61197.1; -; mRNA.  
CC EMBL: AF115311; AAF21970.1; -; mRNA.  
CC EMBL: AF176342; AAK64595.1; -; mRNA.  
CC EMBL: AF251263; AAF98152.1; -; mRNA.  
CC EMBL: AL590119; CAC34872.1; -; mRNA.  
CC EMBL: CR456588; CAG30474.1; -; mRNA.  
CC EMBL: Z97055; CAB09788.1; -; Genomic\_DNA.  
CC EMBL: BC022459; AAH22459.1; -; mRNA.  
CC EMBL: BC028171; AAH28171.1; -; mRNA.  
CC EMBL: BC030665; AAH30665.1; -; mRNA.  
CC HSSP: P50224; 1CJM.  
CC Ensembl: ENSG00000130540; Homo sapiens.  
CC HGNC: HGNC:14903; SULT4A1.  
CC MIM: 608359; gene.  
CC GO: GO:0005829; C:cytosol; NAS.  
CC GO: GO:0008146; F:sulfotransferase activity; NAS.  
CC InterPro: IPR000863; Sulfotransferase.  
CC Pfam: PF00685; Sulfotransferase 1; 1.  
CC ProDom: PD001218; Sulfotransferase; 1.  
CC Alternative splicing: Lipid metabolism; Steroid metabolism;  
CC Transference.  
KW CHAIN 1 284 Sulfotransferase 4A1.  
FT FT /FTID=PRO 0000085167.  
FT NP\_BIND 246 254 PAPS (Potential).  
FT VAESPLIC 248 284 GRVGLWKDIFTVSMNEKEDLVYKQMGKCDLTFDFYL ->  
FT AHCVFARKIFLSW (in isoform 2).  
FT /FTID=VSP 006304.  
FT CONFLICT 55 56 KS -> P (in Ref. 8; AAH30665).  
FT CONFLICT 239 239 N -> S (in Ref. 8; AAH22459).  
SQ SEQUENCE 284 AA; 33085 MW; A6EA6844B66C400B CRC64;  
  
Query Match 100.0%; Score 1530; DB 1; Length 284;  
Best Local Similarity 100.0%; Pred. No. 1.2e-119;  
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MAESEATPTSGEPESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60  
Db 1 MAESEATPTSGEPESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60  
  
QY 61 LQEVVYLSQADPDEIGLWNIDEQLPVLEYQPGLDIIKELTSRLIKSHLPYRFLPSD 120  
Db 61 LQEVVYLSQADPDEIGLWNIDEQLPVLEYQPGLDIIKELTSRLIKSHLPYRFLPSD 120  
  
QY 121 LHNGDSKVIYARNPKDLVSVYQPHRSRLTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180  
Db 121 LHNGDSKVIYARNPKDLVSVYQPHRSRLTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180  
  
QY 181 EFWEHRMDSNVLFLEYEDMHRDLVTMVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCNA 240  
Db 181 EFWEHRMDSNVLFLEYEDMHRDLVTMVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCNA 240  
  
QY 241 EALPVGGRVGLWKDIFTVSMNEKEDLVYKQMGKCDLTFDFYL 284  
Db 241 EALPVGGRVGLWKDIFTVSMNEKEDLVYKQMGKCDLTFDFYL 284  
  
RESULT 2  
Q5R4X3\_PONPY PRELIMINARY; PRT; 284 AA.  
ID Q5R4X3\_PONPY  
AC Q5R4X3\_PONPY  
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License

DT 07-FEB-2006, entry version 7.  
DE Hypothetical protein DKFP459L1114.  
GN Name=DKFP459L1114;  
OS Pongo pygmaeus (Orangutan).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Pongo.  
OX NCBI\_TaxID=9600;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Cortex;  
RG The German cDNA Consortium;  
RA Bloeker H., Boecher M., Brandt P., Mewes H.W., Weil B., Amid C.,  
RA Osanger A., Fobo G., Han M., Wiemann S.;  
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.  
CC  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC  
CC EMBL: CR611116; CAH93193.1; -; mRNA.  
CC GO: GO:0008146; F:sulfotransferase activity; IEA.  
CC InterPro: IPR000863; Sulfotransferase.  
CC Pfam: PF00685; Sulfotransferase 1; 1.  
CC ProDom: PD001218; Sulfotransferase; 1.  
CC Hypothetical protein.  
KW SEQUENCE 284 AA; 33111 MW; 531F6C10A3294B00 CRC64;  
  
Query Match 99.7%; Score 1525; DB 2; Length 284;  
Best Local Similarity 99.6%; Pred. No. 3.2e-119;  
Matches 283; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 MAESEATPTSGEPESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60  
Db 1 MPSEAEATPTSGEPESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60  
  
QY 61 LQEVVYLSQADPDEIGLWNIDEQLPVLEYQPGLDIIKELTSRLIKSHLPYRFLPSD 120  
Db 61 LQEVVYLSQADPDEIGLWNIDEQLPVLEYQPGLDIIKELTSRLIKSHLPYRFLPSD 120  
  
QY 121 LHNGDSKVIYARNPKDLVSVYQPHRSRLTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180  
Db 121 LHNGDSKVIYARNPKDLVSVYQPHRSRLTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180  
  
QY 181 EFWEHRMDSNVLFLEYEDMHRDLVTMVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCNA 240  
Db 181 EFWEHRMDSNVLFLEYEDMHRDLVTMVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCNA 240  
  
QY 241 EALPVGGRVGLWKDIFTVSMNEKEDLVYKQMGKCDLTFDFYL 284  
Db 241 EALPVGGRVGLWKDIFTVSMNEKEDLVYKQMGKCDLTFDFYL 284  
  
RESULT 3  
Q865L7\_RABIT PRELIMINARY; PRT; 284 AA.  
ID Q865L7\_RABIT  
AC Q865L7;  
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.  
DT 01-JUN-2003, sequence version 1.  
DE 07-FEB-2006, entry version 11.  
DE Brain sulfotransferase 4A1.  
GN Name=Sult4a1;  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;  
OC Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA He D., Falany C.N.;  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
CC  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License

```

CC -----
DR EMBL; AY196782; AAO45181.1; -, mRNA.
DR HSPF; P50224; ICM.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR Transf.
SQ SEQUENCE 284 AA; 33107 MW; 782B6899407A090DE CRC64;

      Query Match          98.6%; Score 1508; DB 2; Length 284;
      Best Local Similarity 98.2%; Pred. No. 8.6e-118;
      Matches 279; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAESEAETPTGPFESKYFHFHGVRLPPFCRGKMEIEANFVRPSDWMIVTPKSGTSL 60
DB 1 MAESEAETPTGPFESKYFHFHGVRLPPFCRGKMEIEANFVRPSDWMIVTPKSGTSL 60
QY 61 LOEVVYLYSQADPDPEIGMNIDEQLPVLEYPOPGLDIIKELTSPRLIKSHLPYRFLPSD 120
DB 61 LOEVVYLYSQADPDPEIGMNIDEQLPVLEYPOPGLDIIKELTSPRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYARNPKDLVSVYQPHRSILRTMSYRGTFQEFCCRPMNDKLGYSWFHVVQ 180
DB 121 LHNGDSKVIYARNPKDLVSVYQPHRSILRTMSYRGTFQEFCCRPMNDKLGYSWFHVVQ 180
QY 181 EFWEHRMDSNVLFLKYEDHMRDLVTMVEQLARFLVSCDKAQLEALTEHCHEQVLDQCQNA 240
DB 181 EFWEHRMDSNVLFLKYEDHMRDLVTMVEQLARFLVSCDKAQLEALTEHCHEQVLDQCQNA 240
QY 241 EALPVGRGRVGLWKDIFTVSMNEKEDLVYKQMGKCDLTFDEYL 284
DB 241 EALPVGRGRVGLWKDIFTVSMNEKEDLVYKQMGKCDLTFDEYL 284

RESULT 4
ST4A1_MOUSE
ID ST4A1_MOUSE STANDARD; PRT; 284 AA.
AC P63046; O88872; Q91X55; Q9CWMY7; Q9DC97;
DT 31-AUG-2004, integrated into UniProtKB/Swiss-Prot.
DT 31-AUG-2004, sequence version 1.
DT 07-FEB-2006, entry version 15.
DE Sulfotransferase 4A1 (EC 2.8.2.-) (Brain sulfotransferase-like
DE protein) (msr-STL) (Nervous system sulfotransferase) (NST).
GN Name=St4a1; Synonyms=Sultx3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Murcidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), FUNCTION, AND TISSUE
RP SPECIFICITY.
RC TISSUE=Brain;
RX MEDLINE=22035352; PubMed=12039030; DOI=10.1016/S0378-1119(02)00431-6;
RA Sakakibara Y., Suiko M., Pai T.G., Nakayama T., Takami Y.,
RA Katafuchi J., Liu M.-C.;
RA "Highly conserved mouse and human brain sulfotransferases: molecular
RT cloning, expression, and functional characterization.";
RL Gene 285:35-47(2002).
RN [2]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).
RP STRAIN=C57BL/6J; TISSUE=Embryonic stem cell;
RX PubMed16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Bremner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilmom L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA Banasi M., Baxter L., Belsel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,

```

```

RA di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Grigucinic S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Huminicki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Lluni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuura S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakatsuji H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C., Sheng Y.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugita K., Sultana R., Takenaka Y., Taki K.,
RA Tannoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Walthers S.M., Teasdale R.D., Liu E.T., Bruscia V., Quackenbush J.,
RA Walstead C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Aragawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Wachihi A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RP STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Roak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May catalyze the sulfate conjugation of many drugs,
CC xenobiotic compounds, hormones, and neurotransmitters. Displays
CC activity towards 1-triiodothyronine, thyroxine, estrone, p-
CC nitrophenol, 2-naphthylamine, and 2-naphtol. May have a role in the
CC sulfation of drugs and neurotransmitters in the CNS.
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P63046-1, Q9DC97-1;
CC Sequence=Displayed;
CC Name=2;
CC IsoId=P63046-2, Q9DC97-2;
CC Sequence=VSP_006305;
CC -!- TISSUE SPECIFICITY: Expressed in brain, cerebellum and
CC hypothalamus. Not detected in pancreas, liver, lung, intestine,
CC kidney, uterus, adrenal gland, thymus, spleen, epididymis,
CC testicle, and heart.

```

```
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AF059257; AAC63999.1; -; mRNA.
DR EMBL: AK003034; BAB22522.1; -; mRNA.
DR EMBL: AK010293; BAB26829.1; -; mRNA.
DR EMBL: BC051132; AAH51132.1; -; mRNA.
DR EMBL: BC054757; AAH54757.1; -; mRNA.
DR HSSP: P50224; 1CJM.
DR MGI: MGI:1888971; Sult4a1.
DR GO: GO:0005737; C:cytoplasm; TAS.
DR GO: GO:0008146; F:sulfotransferase activity; IDA.
DR GO: GO:0008790; F:sulfur metabolism; IDA.
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransferase 1; 1.
KW Alternative splicing; Lipid metabolism; Steroid metabolism;
KW Transferase.
FT CHAIN 1 284 Sulfotransferase 4A1.
FT NP_BIND 246 254 /FTID=PRO_0000085168.
FT VARSPLIC 248 284 GRAVLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL ->
FT AHCLEFOTKIALRWRCGRSGSLHCLDLVHVTA (in isoform 2).
FT FTID=VSP_006305.
FT CONFLICT 6 6 A -> R (in Ref. 2; BAB22522).
FT SEQUENCE 284 AA; 33054 MW; FCAE940F7219B6FC CRC64;

Query Match 98.4%; Score 1506; DB 1; Length 284;
Best Local Similarity 97.9%; Pred. No. 1.3e-117; Indels 0; Gaps 0;
Matches 278; Conservative 4; Mismatches 2;

QY 1 MAESEAETPTGPEFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAETPTGPEFESKYFEFHGVRLLPPFCRGKMEETADFPVRPSDVWIVTPKSGTSL 60
QY 61 LOEVVYLVSQGADPDEIGLMNIDQOLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQGADPDEIGLMNIDQOLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTWTSYRGTFQEFCCRPMNDKLGVSWEHVQ 180
Db 121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTWTSYRGTFQEFCCRPMNDKLGVSWEHVQ 180
QY 181 EFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCHQLVDQCCNA 240
QY 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 5
ST4A1_RAT STANDARD; PRT; 284 AA.
AC P63047; O88872; Q91XS5; Q9C9Y7; Q9DC97;
DT 31-AUG-2004, integrated into UniProtKB/Swiss-Prot.
DT 31-AUG-2004, sequence version 1.
DT 07-FEB-2006, entry version 14.
DE Sulfotransferase 4A1 (EC 2.8.2.-) (Brain sulfotransferase-like protein) (rBR-STL) (Nervous system sulfotransferase) (NST).
GN Name=Sult4a1; Synonyms=Sultx3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], TISSUE SPECIFICITY, DEVELOPMENTAL STAGE,
RP AND SUBCELLULAR LOCATION.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
```

```
RX MEDLINE=20164074; PubMed=10698717; DOI=10.1042/0264-6021.3460857;
RA Falany C.N., Xie X., Wang J., Ferrer J., Falany J.L.;
RT "Molecular cloning and expression of novel sulphotransferase-like
RT cDNAs from human and rat brain.";
RL Biochem. J. 346:857-864(2000).
RN [2]
RN NUCLEOTIDE SEQUENCE [MRNA] OF 125-284.
RP TISSUE=Brain;
RC Martin S.C., Farb D.H.;
RT "Molecular identification of a rat nervous system cytoplasmic
RT sulfotransferase, NST.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May catalyze the sulfate conjugation of many drugs,
CC xenobiotic compounds, hormones, and neurotransmitters. Displays
CC activity towards L-triiodothyronine, thyroxine, estrone, p-
CC nitrophenol, 2-naphthylamine, and 2-naphthol. May have a role in the
CC sulfation of drugs and neurotransmitters in the CNS (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- TISSUE SPECIFICITY: Expressed in the brain, not detected in the
CC liver, kidney, spleen, heart, small intestine or testis.
CC -!- DEVELOPMENTAL STAGE: Expressed at low levels in brains of 1-day
CC old animals but increase to adult levels from 7-day old animals
CC and remain at that level in adults.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AF188699; AAF61198.1; -; mRNA.
DR EMBL: AF176343; AAK84596.1; -; mRNA.
DR HSSP: P50224; 1CJM.
DR RGD: 69292; Sult4a1.
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransferase 1; 1.
KW Lipid metabolism; Steroid metabolism; Transferase.
FT CHAIN 1 284 Sulfotransferase 4A1.
FT NP_BIND 246 254 /FTID=PRO_0000085169.
FT SEQUENCE 284 AA; 33054 MW; FCAE940F7219B6FC CRC64;

Query Match 98.4%; Score 1506; DB 1; Length 284;
Best Local Similarity 97.9%; Pred. No. 1.3e-117; Indels 0; Gaps 0;
Matches 278; Conservative 4; Mismatches 2;

QY 1 MAESEAETPTGPEFESKYFEFHGVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSL 60
Db 1 MAESEAETPTGPEFESKYFEFHGVRLLPPFCRGKMEETADFPVRPSDVWIVTPKSGTSL 60
QY 61 LOEVVYLVSQGADPDEIGLMNIDQOLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
Db 61 LOEVVYLVSQGADPDEIGLMNIDQOLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120
QY 121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTWTSYRGTFQEFCCRPMNDKLGVSWEHVQ 180
Db 121 LHNGDSKVIYMARNPKDLVSVYQFHRSLRTWTSYRGTFQEFCCRPMNDKLGVSWEHVQ 180
QY 181 EFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCHQLVDQCCNA 240
Db 181 EFWEHRMDSNVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCHQLVDQCCNA 240
QY 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
Db 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284

RESULT 6
Q3TX5_MOUSE PRELIMINARY; PRT; 284 AA.
AC Q3TX5;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
```

DE Visual cortex cDNA, RIKEN full-length enriched library,  
DE clone.K530039109 product:sulfotransferase family 4A, member 1, full  
DE insert sequence.  
GN Name=Sulf4a1;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Visual cortex;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Methods Enzymol. 303:19-44(1999).  
[2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Visual cortex;  
RX PubMed16141073; DOI=10.1126/science.1112014;  
RA Carninci P., Kasukawa T., Katayama S., Gough J., Hrtch J., Ishikawa N.,  
RA Oyama R., Ravasi S., Lenhard B., Wells C., Kodzius R., Shimokawa K.,  
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,  
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,  
RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,  
RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,  
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,  
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,  
RA Di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,  
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,  
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,  
RA Guerniche S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,  
RA Hill D., Hummel L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,  
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Keiso J., Kitamura H.,  
RA Kitano H., Kollipati G., Krishnan S.P., Kruger A., Kummerfeld S.K.,  
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,  
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,  
RA Matsuura H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,  
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,  
RA Nillesen R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,  
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,  
RA Petkovsky N., Piazza S., Reed J.C., Reid J.F., Ring B.Z., Ringwald M.,  
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,  
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,  
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,  
RA Sperling S., Stupka E., Sugura K., Sultana R., Takenaka Y., Taki K.,  
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,  
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,  
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,  
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,  
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,  
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,  
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,  
RA Kawashima T., Kojima K., Kondo S., Konno H., Nakano K., Nimomiya N.,  
RA Niehio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,  
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,  
RA Hayashizaki Y.;  
RT "The transcriptional landscape of the mammalian genome.";  
RL Science 309:1559-1563(2005).  
[3]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Visual cortex;  
RX PubMed16141073; DOI=10.1126/science.1112009;  
RG RIKEN Genome Exploration Research Group, and Genome Science Group  
RG (Genome Network Core Team) and the FANTOM Consortium;  
RT "Antisense Transcription in the Mammalian Transcriptome";  
RL Science 309:1564-1566(2005).  
[4]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Visual cortex;  
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,

Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
RA Blake J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,  
RA Dalla E., Dragan T.A., Fletcher C.F., Forrest A., Gough J.,  
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,  
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tonita M.,  
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
[5]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Visual cortex;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Iwama M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Stauber F., Suzuki R., Tonita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzatelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsuke S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 403:685-690(2001).  
[6]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Visual cortex;  
RX MEDLINE=20493374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
[7]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Visual cortex;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Konno H., Kiyama J., Nishi K., Kito T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multipipette sequencer.";



```
RL  Genome Res. 10:1757-1771 (2000).
RN  [8]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=C57BL/6J; TISSUE=Visual cortex;
RA  Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayaishida K.,
RA  Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA  Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA  Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA  Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA  Muramatsu M., Hayaishizaki Y.;
RA  Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC  -----
CC  Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC  Distributed under the Creative Commons Attribution-NoDerivs License
CC  -----
DR  EMBL; AK159034; BAE34779.1; -; mRNA.
DR  MGI; MGI:1888971; Sult4a1.
DR  GO; GO:0005737; C:cytoplasm; TAS.
DR  GO; GO:0008146; F:sulfotransferase activity; IDA.
DR  GO; GO:0008146; F:sulfotransferase activity; RCA.
DR  GO; GO:0006790; P:sulfur metabolism; IDA.
DR  InterPro; IPR000863; Sulfotransferase.
DR  Pfam; PF00685; Sulfotransferase 1; 1.
DR  ProDom; PD001218; Sulfotransferase; 1.
KW  Transferase.
SQ  SEQUENCE 284 AA; 33054 MW; FCAE940F7219B6FC CRC64;

Query Match          98.4%; Score 1506; DB 2; Length 284;
Best Local Similarity 97.9%; Pred. No. 1.3e-117;
Matches 278; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY  1 MAESEAETPTGPEFSKYFEHGVRLPPFCRGKMEIEANFVRPSDVWIVTPKSGTSL 60
Db  1 MAESEAETPTGPEFSKYFEHGVRLPPFCRGKMEIEADFFVRPSDVWIVTPKSGTSL 60
QY  61 LOEVVTVLSQGDPPDEIGLWNIDEQLPVLEYQPGGLDIIKELTSPLIKSHLPYRFLPSD 120
Db  61 LOEVVTVLSQGDPPDEIGLWNIDEQLPVLEYQPGGLDIIKELTSPLIKSHLPYRFLPSD 120
QY  121 LHNGDSKVITYMARNPKDLVSVYQFHRSLRTWSYRGTFOFCRRFNMNDKLGSGWFHVQ 180
Db  121 LHNGDSKVITYMARNPKDLVSVYQFHRSLRTWSYRGTFOFCRRFNMNDKLGSGWFHVQ 180
QY  181 EFWEHRMDSNVFLKYEDMHRDLVTNVEQLARFLGVSCDKAQLAEALTECHQLVQCCNA 240
Db  181 EFWEHRMDSNVFLKYEDMHRDLVTNVEQLARFLGVSCDKAQLAEALTECHQLVQCCNA 240
QY  241 EALPVGRGVGLWKDIFTVSMNEKEDLVYKQMGKCDLTFDFYL 284
Db  241 EALPVGRGVGLWKDIFTVSMNEKEDLVYKQMGKCDLTFDFYL 284

RESULT 7
QBQV8 MOUSE PRELIMINARY; PRT; 284 AA.
ID  QBQV8 MOUSE
AC  QBQV8;
DT  01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT  01-MAR-2003, sequence version 1.
DT  07-FEB-2006, entry version 20.
DE  Adult male corpora quadrigemina cDNA, RIKEN full-length enriched
DE  library, clone:B230376P15 product:sulfotransferase family 4A, member
DE  1, full insert sequence.
GN  Name=Sult4a1;
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC  Muridae; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  TISSUE=Corpora quadrigemina; STRAIN=C57BL/6J;
RC  MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RX  Carninci P., Hayaishizaki Y.;
```

```
RT  "High-efficiency full-length cDNA cloning.";
RN  Methods Enzymol. 303:19-44 (1999).
RN  [2]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX  PubMed=16141072; DOI=10.1126/science.1112014;
RX  Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA  Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA  Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA  Davis M.J., Wilming L.G., Aldonis V., Allen J.E.,
RA  Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA  Bansal M.P., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA  Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA  Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA  di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA  Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA  Georgii-Heminger P., Gingeras T.R., Gojobori T., Green R.E.,
RA  Gustincich S., Harbers M., Hayaishi Y., Hensch T.K., Hirokawa N.,
RA  Hill D., Huminecki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA  Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA  Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA  Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA  Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA  Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA  Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
RA  Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA  Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA  Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA  Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA  Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA  Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA  Sperling S., Stupka E., Sugitara K., Sultana R., Takenaka Y., Taki K.,
RA  Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA  Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA  Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA  Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA  Wahleschdt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA  Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA  Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA  Kawashima T., Kojima M., Konno S., Konno K., Nakano K., Ninomiya N.,
RA  Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,
RA  Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA  Hayaishizaki Y.;
RT  "The transcriptional landscape of the mammalian genome.";
RN  Science 309:1559-1563 (2005).
RN  [3]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX  PubMed=16141073; DOI=10.1126/science.1112009;
RG  RIKEN Genome Exploration Research Group, and Genome Science Group
RG  (Genome Network Core Team) and the FANTOM Consortium;
RT  "Antisense Transcription in the Mammalian Transcriptome.";
RN  Science 309:1564-1566 (2005).
RN  [4]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX  MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RX  Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA  Nikaide I., Oeato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA  Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA  Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA  Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA  Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA  Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA  Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA  Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA  Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA  Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA  Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA  Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA  Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA  Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA  Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
```

RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.;  
RA "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573 (2002).  
RN [5]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690 (2001).  
RN [6]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630 (2000).  
RN [7]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Iehikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771 (2000).  
RN [8]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
RA Saito R., Saitoh H., Sakai K., Sakai K., Sakazume N., Sano H.,  
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,

RA Tomaru A., Toyu T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
RL Submitted (JUL-2001) to the ENBL/GenBank/DBJ databases.  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
CC EMBL; AK046368; BAC32692.1; -; mRNA.  
DR HSSP; P50224; 1CJM  
DR Ensemble; ENSMUSG00000018865; Mus musculus.  
DR MGI; MGI:1888971; Sult4a1.  
DR GO; GO:0005737; C:cytoplasm; TAS.  
DR GO; GO:0008146; F:sulfotransferase activity; IDA.  
DR GO; GO:0008146; F:sulfotransferase activity; RCA.  
DR GO; GO:0006790; P:sulfur metabolism; IDA.  
DR InterPro; IPR000863; Sulfotransferase.  
DR Pfam; PF00685; Sulfotransfer\_1; 1.  
DR ProDom; PD001218; Sulfotransferase; 1.  
KW Transferase.  
SQ SEQUENCE 284 AA; 32972 MW; 38D7B556F2B9391E CRC64;  
Query Match 96.9%; Score 1483; DB 2; Length 284;  
Best Local Similarity 96.5%; Pred. No. 1.1e-115;  
Matches 274; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
QY 1 MAESEATPTSPGFEFESKYFEFHGVRLPPFCRGKMEIANPPVRPSDVWIVYPKSGTSL 60  
DB 1 MGSEATPTGTPGFEFESKYFEFHGVRLPPFCRGKMEIANPPVRPSDVWIVYPKSGTSL 60  
QY 61 LQEVYLVSGQADPEIGLNMNIDQLVLEYPQGLDIKELTSPRLIKSHLPYRFLPSD 120  
DB 61 LQEVYLVSGQADPEIGLNMNIDQLVLEYPQGLDIKELTSPRLIKSHLPYRFLPSD 120  
QY 121 LHNQDSKVIYARNPKDLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFEHVQ 180  
DB 121 LHNQDSKVIYARNPKDLVSYQFHRSLRTMSYRGTFQECRRFMNDKLGYSWFEHVQ 180  
QY 181 EFWHRMDSNVLFUKYEDMHRDLVTWVEQLARFVGSCDKAQLALTEHCHQLVDQCNA 240  
DB 181 EFWHRMDANVLFUKYEDMHRDLVTWVEQLARFVGSCDKAQLALTEHCHQLVDQCNA 240  
QY 241 EALPVGGRVGLWKDIFTVSMNKEFDLVYQKMGKCDLTDFDFYL 284  
DB 241 EALPVGGRVGLWKDIFTVSMNKEFDLVYQKMGKCDLTDFDFYL 284  
RESULT 8  
Q68EV3\_XENLA  
ID Q68EV3\_XENLA PRELIMINARY; PRT; 284 AA.  
AC Q68EV3;  
DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.  
DT 11-OCT-2004, sequence version 1.  
DT 07-FEB-2006, entry version 12.  
DE MGC84327 protein.  
GN Name=MGC84327;  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Eye;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.;  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative.";  
RL Dev. Dyn. 225:384-391 (2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Eye;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RA "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Eye;  
RA Klein S., Gerhard D.S.;  
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; BC080097; AAH80097.1; -; mRNA.  
DR GO; GO:0008146; F:sulfotransferase activity; IEA.  
DR InterPro; IPR000863; Sulfotransferase.  
DR Pfam; PF00685; Sulfotransferase 1; 1.  
DR ProDom; PD001218; Sulfotransferase; 1.  
SQ SEQUENCE 284 AA; 33211 MW; 9344EE54A62729D3 CRC64;  
  
Query Match 90.5%; Score 1385; DB 2; Length 284;  
Best Local Similarity 89.1%; Pred. No. 1.7e-107;  
Matches 253; Conservative 18; Mismatches 13; Indels 0; Gaps 0;  
  
QY 1 MAESEAETPTGCFESKYFEHGVRLPPFCRGKMEETANFPVRPSDVIWVTPKSGTSL 60  
DB 1 MAESEAETPTGCFESKYFEHGVRLPPFCRGKMEESDFFVRENDIWIWTPKSGTSL 60  
  
QY 61 LOEVVYLSQAGDPDEIGLMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120  
DB 61 LOEVVYLSQAGDPDEIGLMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120  
  
QY 121 LHNGSKVIYMARPKDLVSVYQFHRSLRTWTSYRGTFQECRRFMNDKLGYSWFHVQ 180  
DB 121 LHNGSKVIYMARPKDLVSVYQFHRSLRTWTSYRGTFQECRRFMNDKLGYSWFHVQ 180  
  
QY 181 EFWHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVQCCNA 240  
DB 181 EFWHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVQCCNA 240  
  
QY 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYQKMGKCDLTFDFYL 284  
DB 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYQKMGKCDLTFDFYL 284  
  
RESULT 9  
Q52KW3 XENLA PRELIMINARY; PRT; 284 AA.  
ID Q52KW3 XENLA  
AC Q52KW3;  
DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.  
DT 24-MAY-2005, sequence version 1.  
DE 07-FEB-2006, entry version 4.  
DE Hypothetical protein.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus; Xenopus.  
OX NCBI\_TaxID=8355;

RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RA "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Eye;  
RA Klein S., Gerhard D.S.;  
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; BC094166; AAH94166.1; -; mRNA.  
DR GO; GO:0008146; F:sulfotransferase activity; IEA.  
DR InterPro; IPR000863; Sulfotransferase.  
DR Pfam; PF00685; Sulfotransferase 1; 1.  
DR ProDom; PD001218; Sulfotransferase; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 284 AA; 33242 MW; FFA81B7E9A8BEAA9 CRC64;  
  
Query Match 90.3%; Score 1382; DB 2; Length 284;  
Best Local Similarity 89.1%; Pred. No. 3.1e-107;  
Matches 253; Conservative 18; Mismatches 13; Indels 0; Gaps 0;  
  
QY 1 MAESEAETPTGCFESKYFEHGVRLPPFCRGKMEETANFPVRPSDVIWVTPKSGTSL 60  
DB 1 MAESEAETPTGCFESKYFEHGVRLPPFCRGKMEESDFFVRENDIWIWTPKSGTSL 60  
  
QY 61 LOEVVYLSQAGDPDEIGLMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120  
DB 61 LOEVVYLSQAGDPDEIGLMNIDQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSD 120  
  
QY 121 LHNGSKVIYMARPKDLVSVYQFHRSLRTWTSYRGTFQECRRFMNDKLGYSWFHVQ 180  
DB 121 LHNGSKVIYMARPKDLVSVYQFHRSLRTWTSYRGTFQECRRFMNDKLGYSWFHVQ 180  
  
QY 181 EFWHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVQCCNA 240  
DB 181 EFWHRMDSNVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVQCCNA 240  
  
QY 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYQKMGKCDLTFDFYL 284  
DB 241 EALPVGGRVGLWKDIFTVSMNEKFDLVYQKMGKCDLTFDFYL 284

```

RESULT 10
Q4SRF8_TETNG
ID Q4SRF8_TETNG PRELIMINARY; PRT; 284 AA.
AC Q4SRF8;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Chromosome undetermined SCAR14526, whole genome shotgun sequence.
GN ORFNames=GSTENG00013931001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jallion O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bounneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Bienmont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.,
RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CAAE01014526; CAF96774.1; -; Genomic DNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
SQ SEQUENCE 284 AA; 32894 MW; 414AB27ED7D8DF9A6 CRC64;

Query Match 85.8%; Score 1313; DB 2; Length 284;
Best Local Similarity 85.6%; Pred. No. 1.8e-101;
Matches 243; Conservative 21; Mismatches 20; Indels 0; Gaps 0;

QY 1 MAESAETPTSTGPEESKYFEPHGVRLPPFCRGKMEETANFPVRPSDWIVTYPKSGTSL 60
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MAAREADTPSTIEPESKYFEPDGVRLPPFCRGKMEETAGFSLRSSDIWIVTYPKSGTSL 60

QY 61 LOEVVYVLSQADPDDEIGLWNIDEOLPVLEYQPGGLDIIKELTSPRLIKSHLPYRFLPSD 120
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 LOEIVLVLSQADPDDEIGLWNIDEOLPVLEYQPGGLDIIKELTSPRLIKSHLPYRFLPTA 120

QY 121 LHNGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 LHNGEAKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVQ 180

QY 161 EFWEHRMDSNVLFKYEDMHRDLVTWVQEARFLVGSCKAQLAEALTECHQLVDVCCNA 240
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
161 EFWEHRMDSNVLFKYEDMHRDLVTWVQEARFLVGSCKAQLAEALTECHQLVDVCCNS 240
161 EFWEHRMDSNVLFKYEDMHRDLVTWVQEARFLVGSCKAQLAEALTECHQLVDVCCNS 240

```

```

QY 241 EALPVGRGRVGLWKDIPTVSMNEKFDLVYKQKMGKCDLTDFYL 284
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 EALSVCRGRVGLWKDVFTVSMNDKFDIYRQKMGKCDLTDFDGL 284

RESULT 11
Q5RCA8_PONPY
ID Q5RCA8_PONPY PRELIMINARY; PRT; 218 AA.
AC Q5RCA8;
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 21-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE Hypothetical protein DKFZp45910535.
GN Name=DKFZp45910535;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Cortex;
RG The German cDNA Consortium;
RA Koehler K., Beyer A., Mewes H.W., Well B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CR858370; CAH90599.1; -; mRNA.
DR SMR; O5RCA8; 35-218.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Hypothetical protein.
SQ SEQUENCE 218 AA; 25328 MW; 87B9496A82B7B878 CRC64;

Query Match 65.0%; Score 995; DB 2; Length 218;
Best Local Similarity 95.9%; Pred. No. 5.7e-75;
Matches 185; Conservative 0; Mismatches 2; Indels 6; Gaps 1;

QY 92 POPGLDIIKELTSPRLIKSHLPYRFLPSDLHNGSKVIYARNPKDLVSVYQFHRSLRT 151
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
32 PSP-----ELTSPRLIKSHLPYRFPSPDLHNGSKVIYARNPKDLVSVYQFHRSLRT 85

QY 152 MSYRGTFQFCRRFMNDKLGYSWFHVQEFWEHRMDSNVLFKYEDMHRDLVTWVQEA 211
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
86 MSYRGTFQFCRRFMNDKLGYSWFHVQEFWEHRMDSNVLFKYEDMHRDLVTWVQEA 145

QY 212 RFLVGSCKAQLAEALTECHQLVDVCCNAEALPVGRGRVGLWKDIPTVSMNEKFDLVYKQ 271
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
146 RFLVGSCKAQLAEALTECHQLVDVCCNAEALPVGRGRVGLWKDIPTVSMNEKFDLVYKQ 205

QY 272 KMGKCDLTDFYL 284
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
206 KMGKCDLTDFYL 218

RESULT 12
Q9R1S5_MOUSE
ID Q9R1S5_MOUSE PRELIMINARY; PRT; 298 AA.
AC Q9R1S5;
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 18.
DE Arylsulfotransferase STIA4.
GN Name=Sult1a1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.

```





RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";   
 RL Methods Enzymol. 303:19-44(1999).  
 [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Gall bladder;  
 RX PubMed=16141072; DOI=10.1126/science.1112014;  
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,  
 RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,  
 RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,  
 RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,  
 RA Ambesi-Impombato A., Aweller R., Aturaliya R.N.,  
 RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,  
 RA Chiu K.P., Chowdhury V., Christoffels A., Clutterbuck D.R.,  
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,  
 RA Di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,  
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,  
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,  
 RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,  
 RA Hill D., Huminek L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,  
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,  
 RA Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,  
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,  
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,  
 RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,  
 RA Mottagui-Fabriz S., Mulder N., Nakano N., Nakaguchi H., Ng P.,  
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,  
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,  
 RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,  
 RA Roost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,  
 RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,  
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,  
 RA Sperling S., Stupka E., Sugtara K., Sultana R., Takenaka Y., Taki K.,  
 RA Tannoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,  
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,  
 RA Yamanihi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,  
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,  
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,  
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,  
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,  
 RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Nimomiya N.,  
 RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,  
 RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,  
 RA Hayashizaki Y.;  
 RT "The transcriptional landscape of the mammalian genome.";   
 RL Science 309:1559-1563(2005).  
 [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Gall bladder;  
 RX PubMed=16141073; DOI=10.1126/science.1112009;  
 RG RIKEN Genome Exploration Research Group, and Genome Science Group  
 (Genome Network Core Team) and the FANTOM Consortium;  
 RT "Antisense Transcription in the Mammalian Transcriptome.";   
 RL Science 309:1564-1566(2005).  
 [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Gall bladder;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaudo I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hall D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagaishima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,

RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Seto K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs.";   
 RL Nature 420:563-573(2002).  
 [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Gall bladder;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Mateu Y., Nikaudo I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";   
 RL Nature 409:685-690(2001).  
 [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Gall bladder;  
 RX MEDLINE=20499174; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes.";   
 RL Genome Res. 10:1617-1630(2000).  
 [7]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Gall bladder;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multiplexed capillary sequencer.";   
 RL Genome Res. 10:1757-1771(2000).  
 [8]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Gall bladder;  
 RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,  
 RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,  
 RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Nimomiya N.,  
 RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,  
 RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,  
 RA Muramatsu M., Hayashizaki Y.;  
 RL Submitted (MAR-2004) to the ENBL/GenBank/DBJ databases.  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>



```
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AK144336; BAE25838.1; -; mRNA.
DR MGI; MGI:1916333; Sult1c2.
DR GO; GO:0005764; C:lysosome; RCA.
DR GO; GO:0008146; F:sulfotransferase activity; RCA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 296 AA; 34953 MW; D2B1769ACA2BEA42 CRC64;

Query Match          31.1%; Score 476; DB 2; Length 296;
Best Local Similarity 35.3%; Pred No. 2.5e-31;
Matches 102; Conservative 58; Mismatches 107; Indels 22; Gaps 5;

QY 16 ESKYFEFHGVRLPPFCRGKMEBIANFPVRPSDVWTVTPKSGTSLQEVYLVSGQADPD 75
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 10 QTKLEVAGIPLQFTVDNWRQIQTFEAKPDDLICTYPKSGTTWQIIVDMIEQNGDVE 69
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 76 EIGLMNIDEQLPVLEY----POP-GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYM 131
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 70 KCRRTIIQHRHFIEWARPPQPSGVDKANEMAPRILRTHLPTQLLPFSFWTNCKFLYV 129
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 132 ARNPXDLVSYVYQFHRSLRTMSYRGTFQFCRRFMNDKLVGSGFWHFQGEFWEHRMDSNV 191
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 130 ARNAKDCWVSYYTHFRMGQVLPPEPTWDEYFTTFINGQVSGSWFDHVGKWEIRDKYQI 189
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 192 LFLKYEDMHRDLVTWVEQLARFLGVS-----CDKAQLEALTEHCHQLVDQCCNABALP-- 244
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 190 LFLFYEDMKRNPKEBIQKVMQPMGNLDEDDVVDKIVLETSPKMKKE--NPMTNRSTAPKS 247
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 245 -----VGRGRVLGKDIITVSMNEKFDLVYKQKMGKCDLTFDFYL 284
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 248 ILDQISPPMRKGTVDGWNKNHFTVAQNERFDEIYKQKMGRTSLNFSMEL 296
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 16
ST1A3_HUMAN
ID ST1A3_HUMAN STANDARD; PRT; 295 AA.
AC P50224; O95603; Q6ZMJ5;
DC 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-1996, sequence version 1.
DT 21-FEB-2006, entry version 55.
DE Monamine-sulfating phenol sulfotransferase (EC 2.8.2.1) (Aryl
DE sulfotransferase 1A3) (Sulfotransferase, monamine-preferring) (M-PST)
DE (Thermolabile phenol sulfotransferase) (TL-PST) (Placental estrogen
DE sulfotransferase) (Catecholamine-sulfating phenol sulfotransferase)
DE (HAGT3).
GN Name=ST1A3; Synonyms=STM;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=93371391; PubMed=8363592;
RA Zhu X., Veronese M.E., Bernard C.C., Sansom L.N., McManus M.E.;
RT "Identification of two human brain aryl sulfotransferase cDNAs.";
RL Biochem. Biophys. Res. Commun. 195:120-127(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Placenta;
RX MEDLINE=94244843; PubMed=8187949; DOI=10.1016/0303-7207(94)90159-7;
RA Bernier F., Lopez-Solache I., Labrie F., Luu-The V.;
RT "Cloning and expression of cDNA encoding human placental estrogen
RT sulfotransferase.";
RL Mol. Cell. Endocrinol. 99:R11-R15(1994).
RN [3]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RX MEDLINE=95100963; PubMed=7802665;
```

```
RA Dooley T.P., Probst P., Munroe P.B., Mole S.E., Liu Z., Doggett N.A.;
RT "Genomic organization and DNA sequence of the human catecholamine-
RT sulfating phenol sulfotransferase gene (STM).";
RL Biochem. Biophys. Res. Commun. 205:1325-1332(1994).
RN [4]
RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 84-101.
RC TISSUE=Liver;
RX MEDLINE=94161723; PubMed=8117269;
RA Wood T.C., Aksoy I.A., Aksoy S., Weinshilboum R.M.; cDNA cloning,
RA "Human liver thermolabile phenol sulfotransferase: cDNA cloning,
RT expression and characterization.";
RL Biochem. Biophys. Res. Commun. 198:1119-1127(1994).
RN [5]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Platelet;
RX MEDLINE=95209696; PubMed=7695637;
RA Aksoy I.A., Weinshilboum R.M.;
RT "Human thermolabile phenol sulfotransferase gene (STM): molecular
RT cloning and structural characterization.";
RL Biochem. Biophys. Res. Commun. 208:786-795(1995).
RN [6]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Blood;
RX MEDLINE=95209704; PubMed=7695643;
RA Jones A.L., Hagen M., Coughtrie M.W.H., Roberts R.C., Glatt H.;
RT "Human platelet phenol sulfotransferase: cDNA cloning, stable
RT expression in V79 cells and identification of a novel allelic variant
RT of the phenol-sulfating form.";
RL Biochem. Biophys. Res. Commun. 208:855-862(1995).
RN [7]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Leukocyte;
RX MEDLINE=95050600; PubMed=7961757;
RA Bernier F., Labrie G., Labrie F., Luu-The V.;
RT "Structure of human estrogen and aryl sulfotransferase gene. Two mRNA
RT species issued from a single gene.";
RL J. Biol. Chem. 269:28200-28205(1994).
RN [8]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Liver;
RA Gaedigk A., Grant D.M.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [9]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoa K., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshihawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigetake K., Senba T.,
RA Matsumura K., Kajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita K.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
```



"Complete sequencing and characterization of 21,243 full-length human cDNAs.";  
Nat. Genet. 36:40-45(2004).  
[10]  
TISSUE=Brain;  
MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
Brownstein M.J., Ustin T.B., Tohyuki S., Carninci P., Prange C.,  
Raha S.S., Lioquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
Boasak S.A., McEwan P.C., McKernan K.J., Malek J.A., Gunaratne P.H.,  
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Whiting J., Helton E., Kretzeman M., Madan A., Rodrigues S., Sanchez A.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,  
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[11]  
TISSUE=Brain;  
MEDLINE=94379981; PubMed=8093002;  
Veronese M.E., Burgess W., Zhu X., McManus M.E.;  
"Functional characterization of two human sulphotransferase cDNAs that encode monoamine- and phenol-sulphating forms of phenol sulphotransferase: substrate kinetics, thermal-stability and inhibitor-sensitivity studies.";  
Biochem. J. 302:497-502(1994).  
[12]  
CHARACTERIZATION.  
MEDLINE=94379981; PubMed=8093002;  
Veronese M.E., Burgess W., Zhu X., McManus M.E.;  
"Functional characterization of two human sulphotransferase cDNAs that encode monoamine- and phenol-sulphating forms of phenol sulphotransferase: substrate kinetics, thermal-stability and inhibitor-sensitivity studies.";  
Biochem. J. 302:497-502(1994).  
[13]  
X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
TISSUE=Brain;  
MEDLINE=20013065; PubMed=10543947; DOI=10.1006/jmbi.1999.3153;  
Bidwell L.M., McManus M.E., Gaedigk A., Kakuta Y., Negishi M., Pedersen L., Martin J.L.;  
"Crystal structure of human catecholamine sulphotransferase.";  
J. Mol. Biol. 293:521-530(1999).  
-!- FUNCTION: Catalyzes the sulfate conjugation of phenolic monoamines (neurotransmitters such as dopamine, norepinephrine and serotonin) and phenolic and cathelic drugs.  
-!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + a phenol = adenosine 3',5'-bisphosphate + an aryl sulfate.  
-!- SUBUNIT: Homodimer.  
-!- SUBCELLULAR LOCATION: Cytoplasm.  
-!- ALTERNATIVE PRODUCTS:  
Event=Alternative splicing; Named isoforms=2;  
Name=1;  
IsoId=P50224-1; Sequence=Displayed;  
Name=2;  
IsoId=P50224-2; Sequence=VSP\_012326, VSP\_012327;  
Note=No experimental confirmation available;  
-!- TISSUE SPECIFICITY: Liver, colon, kidney, lung, brain, spleen, small intestine, placenta and leukocyte.  
-!- PTM: The N-terminus is blocked.  
-!- SIMILARITY: Belongs to the sulphotransferase family.

```

-----
CC EMBL; L19956; AAA02943.1; -; mRNA.
DR EMBL; L25275; AAA36523.1; -; mRNA.
DR EMBL; U08032; AAA17723.1; -; mRNA.
DR EMBL; U20499; AAA64490.1; -; Genomic_DNA.
DR EMBL; X84653; CAA59146.1; -; mRNA.
DR EMBL; L34160; -; NOT_ANNOTATED_CDS; Genomic_DNA.
DR EMBL; U37686; AAA86536.1; -; Genomic_DNA.
DR EMBL; U34199; AAC99987.1; -; mRNA.
DR EMBL; AK122733; BAC85507.1; -; mRNA.
DR EMBL; BC014471; AAH14471.1; -; mRNA.
DR EMBL; BC078144; AAH78144.1; -; mRNA.
DR EMBL; U08099; AAA82126.1; -; Genomic_DNA.
DR PIR; A55451; A55451.
DR PDB; 1CJW; X-ray; A=1-295.
DR PDB; 2A3R; X-ray; A/B=1-295.
DR OGP; P50224; -.
DR H-InVDB; HIX0012928; -.
DR HGNC; HGNC:11455; SULTIA3.
DR MIM; 600641; gene.

Query Match          30.8%; Score 472; DB 1; Length 295;
Best Local Similarity 35.8%; Pred. No. 5 4e-31;
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;

QY 24 GVRLPPFCRGKMEBIANFPVRSDVMIVTPKSGTSLLEQVVVLVSQGADPDEIGLMNID 83
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
DB 17 GVPLIKYFAEALGFLQSFOARPDLLINTYPKSGTTWSQILDMIYGCGDLEKCNRAPIY 76

QY 84 EQLPVLEYPOP---GLDIIKETLSRLIASHLYPYRLPSDLHGNGSKVIYMARNPKDLV 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 77 VRVPFLEVNDPGSPSGLLETUKDTPPRLIKSHILPLLQPOTLLDQKYGVVYVARNPXDVA 136

QY 140 VSYQQPHRSRLTWSYRGTFQFCRCRFMNDKLGVGSMFEHVQBEWEHRMDSNVLFKYEDM 199
   |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 137 VSYYHFHRMEKAHPCTGWDSFLEKFWAGSVGSYQHVOEWELSRTHPVLYLFYEDM 196

QY 200 HRDLVTMVEQLARFLGVSCDKAQLEALTEHC-----HQLVDCCNAEA 242
   ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
DB 197 KENPKREIQILEFVGRSLPEETMDFMVQHTSFKEMKKNPWNTVTPQELMDHSIS-- 253

QY 243 LP-VGRGRVGLWKDIFTVSNNEKFDLVYKQMGKCDLTFF 280
   : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 254 -PFMRKGMGDWKTTFTVAQNERFPDADYAEEKMAGCSLSF 291

RESULT 17
STKL_RAT ID-STKL_RAT STANDARD; PRT; 296 AA.
AC Q9WUW8;
DT 01-DEC-2000, integrated into UniProtKB/Swiss-Prot.
DT 01-NOV-1999, sequence version 1.
DT 21-FEB-2006, entry version 32.
DE Sulfotransferase K1 (EC 2.8.2.-) (*SULT1C2).
GN Name=Sultk1; Synonyms=Sult1c2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
ON [1]
RR NUCLEOTIDE SEQUENCE [MRNA].
RP TTSUSE=Kidney;
RX MEDLINE=20329235; PubMed=10872834; DOI=10.1006/bbrc.2000.2744;
RA Xiangrong L., Joehnk C., Hartmann D., Schestag F., Kroemer W.,
RA Gieselmann V.;
RT "Enzymatic properties, tissue-specific expression, and lysosomal
RT location of two highly homologous rat SULT1C2 sulfotransferases.";
RL Biochem. Biophys. Res. Commun. 272:242-250(2000).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of many drugs,
CC xenobiotic compounds, hormones, and neurotransmitters.
CC -!- SUBCELLULAR LOCATION: Lysosome.
CC -!- TISSUE SPECIFICITY: Highly expressed in kidney and at lower levels
```

- TISSUE SPECIFICITY: Highly expressed in kidney and at lower levels
- SUBCELLULAR LOCATION: Lysosome.
- MEMBRANE COMPONENT: This non-membrane protein.

CC in stomach and liver. More specifically found in the epithelia of  
 CC proximal tubules of the kidney, of the bile duct, of the gastric  
 CC mucosa, and in hepatocytes.

CC -!- SIMILARITY: Belongs to the sulfotransferase family.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----  
 CC EMBL: AJ238391; CAB41460.1; -; mRNA.  
 CC PIR: JC7282; JC7282.

CC HSSP: P50224; 1CQM.

CC DR Ensembl: ENSRNOG0000031833; Rattus norvegicus.

CC RGD: 621064; Sulf1c2.

CC InterPro: IPR000863; Sulfotransferase.

CC Pfam: PF0685; Sulfotransferase\_1.

CC DR ProDom: PD01218; Sulfotransferase; 1.

CC KW Lysosome; Transferase.

CC FT CHAIN 1 296

CC FT Sulfotransferase K1.

CC FT /FTID=PRO 0000085135.

CC FT NP\_BIND 49 54

CC FT PAPS (By similarity).

CC FT NP\_BIND 131 139

CC FT PAPS (By similarity).

CC FT NP\_BIND 194 230

CC FT PAPS (By similarity).

CC FT NP\_BIND 258 260

CC FT PAPS (By similarity).

CC FT ACT\_SITE 109 109

CC FT Proton acceptor (By similarity).

CC FT SEQUENCE 296 AA; 34699 MW; 4A82FAD740EBABF5 CRC64;

CC SQ

CC Query Match 30.5%; Score 467; DB 1; Length 296;

CC Best Local Similarity 35.1%; Pred. No. 1.4e-30;

CC Matches 100; Conservative 58; Mismatches 105; Indels 22; Gaps 5;

CC QY 16 ESKYFEFGVRLPPFCRCGMEEIANFPVRPSDVWIVTPKSGTSLLOEVVYLVSGADPD 75

CC Db 10 QTKLEVAGIPLQAPTVDNWSQIQTFKAKPDDLICTPKSGTTWQIVDMIEQNGDVE 69

CC QY 76 EIGLMNIDEQLPVLEY---POP-GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVYIM 131

CC Db 70 KQRTIIQHRHFFIEWARPPQSGVDKANAMPAPRILRTHLPTQLLPSPFWNNCKFLYV 129

CC QY 132 ARNPKDLVVSYVQFHRSLRTMSYRGTFQFCFRFNDKLGYSWFHVEHFMDSNV 191

CC Db 130 ARNADCMVSYVYHFYRMSQVLDPDGTWNEYPFTFNGKVSWSGDFHVGKWEIRDRYQI 189

CC QY 192 LFLKYEDMHRDLVTMVEQLARFLGVS-----CDKAQLEALTEHCHQLVDQCCNAEALP-- 244

CC Db 190 LFLFYEDVKRDPKREIQKVMQFMGNLDEEVVDKIVLETSEFKMKE--NPMTNRSTVPKS 247

CC QY 245 -----VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280

CC Db 248 VLDQISPPMRKGTGDMKNHFTVAQNDRFDEIYKQKMGTSLSNF 292

CC RESULT 18

CC Q3ZAV3 RAT

CC ID Q3ZAV3 RAT PRELIMINARY; PRT; 296 AA.

CC AC Q3ZAV3

CC DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.

CC DT 07-FEB-2006, sequence version 1.

CC DE Sulfotransferase family, cytosolic, 1C, member 2.

CC GN Name=Sulf1c2;

CC OS Rattus norvegicus (Rat).

CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

CC OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

CC OC Muridae; Murinae; Rattus.

CC OX NCBI\_TaxID=10116;

CC RN [1]

CC RP NUCLEOTIDE SEQUENCE.

CC RC TISSUE=Kidney.

CC RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

CC RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,

CC RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

CC RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

CC RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

CC RA

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RP [2]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Kidney.

RG NIH MGC Project.

RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

CC EMBL: BC103636; AA103637.1; -; mRNA.

CC GO: GO:0008146; F:sulfotransferase activity; IEA.

CC GO: GO:0016740; P:transferase activity; IEA.

CC Transferase.

CC SEQUENCE 296 AA; 34699 MW; 4A82FAD740EBABF5 CRC64;

CC Query Match 30.5%; Score 467; DB 2; Length 296;

CC Best Local Similarity 35.1%; Pred. No. 1.4e-30;

CC Matches 100; Conservative 58; Mismatches 105; Indels 22; Gaps 5;

CC QY 16 ESKYFEFGVRLPPFCRCGMEEIANFPVRPSDVWIVTPKSGTSLLOEVVYLVSGADPD 75

CC Db 10 QTKLEVAGIPLQAPTVDNWSQIQTFKAKPDDLICTPKSGTTWQIVDMIEQNGDVE 69

CC QY 76 EIGLMNIDEQLPVLEY---POP-GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVYIM 131

CC Db 70 KQRTIIQHRHFFIEWARPPQSGVDKANAMPAPRILRTHLPTQLLPSPFWNNCKFLYV 129

CC QY 132 ARNPKDLVVSYVQFHRSLRTMSYRGTFQFCFRFNDKLGYSWFHVEHFMDSNV 191

CC Db 130 ARNADCMVSYVYHFYRMSQVLDPDGTWNEYPFTFNGKVSWSGDFHVGKWEIRDRYQI 189

CC QY 192 LFLKYEDMHRDLVTMVEQLARFLGVS-----CDKAQLEALTEHCHQLVDQCCNAEALP-- 244

CC Db 190 LFLFYEDVKRDPKREIQKVMQFMGNLDEEVVDKIVLETSEFKMKE--NPMTNRSTVPKS 247

CC QY 245 -----VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280

CC Db 248 VLDQISPPMRKGTGDMKNHFTVAQNDRFDEIYKQKMGTSLSNF 292

CC RESULT 19

CC Q642G8 RAT

CC ID Q642G8 RAT PRELIMINARY; PRT; 296 AA.

CC AC Q642G8

CC DT 25-OCT-2004, integrated into UniProtKB/TrEMBL.

CC DT 25-OCT-2004, sequence version 1.

CC DT 07-FEB-2006, entry version 11.

CC DE Sulfotransferase family, cytosolic, 1C, member 1 (predicted).

CC GN Name=Sulf1c1\_predicted;

CC OS Rattus norvegicus (Rat).

CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

CC OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

CC OC Muridae; Murinae; Rattus.

CC OX NCBI\_TaxID=10116;

CC RN [1]

CC RP NUCLEOTIDE SEQUENCE.

CC RC TISSUE=Kidney.

```

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy P.H.,
RA Bhat N.K., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RG NIH MGC Project;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC081691; AAH81691.1; -; mRNA.
DR SMR; Q64268; 12-296.
DR GO; ENSRNOG0000031833; Rattus norvegicus.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
SQ SEQUENCE 296 AA; 34859 MW; FB14A27E39EE847F CRC64;

Query Match 30.5%; Score 466; DB 2; Length 296;
Best Local Similarity 35.4%; Pred. No. 1.7e-30;
Matches 101; Conservative 56; Mismatches 106; Indels 22; Gaps 5;

Qy 16 ESKYFEFHGVRLPPFCRCGMBEIANFPVRPSDVIVTPYKSGTSLLOEVLVLSQADPD 75
Db 10 QTKLKEVAGIPLMDPTVNNWSQIQTFKAKPDDLICTYPKSGTTWQIVNMIEQNGDVE 69

Qy 76 BIGLNMNIDEOLPVLEY---POP-GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYM 131
Db 70 KCQRTIIQHRHPVIEWARPPQPSGVVDKANAMPAPRILRTHLPQLPPSPFTNNCKYLYV 129

Qy 132 ARNPKDLVSVYQFHRSLRTWSYRGTFQFCRRFMNDKLGYSWFHVEQFEWHRMDSNV 191
Db 130 ARNAKDCMVSYHFYRMCQVLNPGTWNYPETFTFNGKVSNGSFDHVKGWMEIRDRYQI 189

Qy 192 LFLKYEDMHRDLVTWVEOLARFLGVS-----CDKAQLEALTEHCHQVDCQCNABALP-- 244
Db 190 LFLFYEDMKRDPKREIQKVMQPMGKNLDEEVVDKIVLETSPFEKMKD--NPLTNFTSTPKT 247

Qy 245 -----VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 248 IMDQISPFMRKGIWGDWKNHFTVAQNERFDEIYEQKLDGTSLNF 292

RESULT 20
Q569D0 RAT PRELIMINARY; PRT; 296 AA.
AC Q569D0
DT 10-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 10-MAY-2005, sequence version 1.
DI 07-FEB-2006, entry version 5.

DE DE
GN GN
OS OS
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy P.H.,
RA Bhat N.K., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RG NIH MGC Project;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC092564; AAH92564.1; -; mRNA.
DR SMR; Q569D0; 12-296.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
SQ SEQUENCE 296 AA; 34875 MW; BB0D65640B8332A3 CRC64;

Query Match 30.4%; Score 465; DB 2; Length 296;
Best Local Similarity 35.3%; Pred. No. 2.1e-30;
Matches 101; Conservative 59; Mismatches 102; Indels 24; Gaps 5;

Qy 16 ESKYFEFHGVRLPPFCRCGMBEIANFPVRPSDVIVTPYKSGTSLLOEVLVLSQADPD 75
Db 10 QTKLKEVAGIPLMDPTVNNWSQIQTFKAKPDDLICTYPKSGTTWQIVNMIEQNGDVE 69

Qy 76 BIGLNMNIDEOLPVLEY---POP-GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYM 131
Db 70 KCQRTIIQHRHPVIEWARPPQPSGVVDKANAMPAPRILRTHLPQLPPSPFTNNCKYLYV 129

Qy 132 ARNPKDLVSVYQFHRSLRTWSYRGTFQFCRRFMNDKLGYSWFHVEQFEWHRMDSNV 191
Db 130 ARNAKDCMVSYHFYRMCQVLNPGTWNYPETFTFNGKVSNGSFDHVKGWMEIRDRYQI 189

Qy 192 LFLKYEDMHRDLVTWVEOLARFLGVS-----CDKAQLEALTEHCHQVDCQCNABALP-- 244
Db 190 LFLFYEDMKRDPKREIQKVMQPMGKNLDEEVVDKIVLETSPFEKMKD--NPLTNFTSTPKT 247

Qy 245 -----VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 248 IMDQISPFMRKGIWGDWKNHFTVAQNERFDEIYEQKLDGTSLNF 292
```

```
RESULT 21
Q2TAB3 HUMAN
ID Q2TAB3 HUMAN PRELIMINARY; PRT; 295 AA.
AC Q2TAB3
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DE 07-FEB-2006, entry version 2.
DE Sulfotransferase family, cytosolic, 1A, phenol-preferring, member 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.C., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RG NIH MGC Project;
RL Submitted (DEC-2005) to the EMBL/GenBank/DBSJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; BC111011; AA111012.1; -; mRNA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR Transferase.
KW SEQUENCE 295 AA; 34325 MW; ECDDEC18DBE80D56 CRC64;

Query Match 30.3%; Score 464; DB 2; Length 295;
Best Local Similarity 35.5%; Pred. No. 2.5e-30;
Matches 99; Conservative 52; Mismatches 102; Indels 26; Gaps 4;

QY 24 GVRLLPFCRCGRKMEIEANFVRPSDVMTVTPKSGTSLLEQVYLVYVLSQADPDDEIGLMNID 83
DB 17 GVPLIKYFAEALGPLQSQFQARPDLLINTPKSGITWTWSQILDMIYQGDLEKCNRAPIY 76

QY 84 EQLPVLVEYPPQ-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDLV 139
DB 77 VRVPLEVNDPGEPSGLLEKDTTPPRLIKSHLPALPQLTLDQKVKVYVYVARNPKDYA 136

QY 140 VSYVYQFHRSLRTMSYRGTFQFCRCRPMNDKLGYSWFHFVQEFWEHRMDSNVLFKLYEDM 199
DB 137 VSYTHFHRMEKAPPEPGTWDSEFLKFMAGEVSYNSWYHQVQEWELSRTHPVLYLYFYEDM 196

QY 200 HRDLVTMVEQLARFLGVSCDKAQLBALTEHC-----HQLVDQCCNABA 242
DB 197 KENPKREIQKILEFVGRSLPEETDMFVQHTSFKEMKKNPMTNTVTPQELMDHSIS--- 253
```

```
QY 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 254 -PFMRKGMAGDWKTKTFTVAQNERFDADYAEKMGCSLSF 291

RESULT 22
Q58CV8 BOVIN
ID Q58CV8 BOVIN PRELIMINARY; PRT; 298 AA.
AC Q58CV8;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DE 07-FEB-2006, entry version 5.
DE Sulfotransferase family, cytosolic, 1C, member 2.
GN Name=SULT1C2;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Pooled;
RX MEDLINE=21180013; PubMed=11282978; DOI=10.1101/gr.170101;
RA Smith T.P.L., Grosse W.M., Freking B.A., Roberts A.J., Stone R.T.,
RA Casas E., Wray J.E., White J., Cho J., Fahrenkrug S.C., Bennett G.L.,
RA Heaton M.P., Laegreid W.W., Rohrer G.A., Chitko-McKown C.G.,
RA Perteira G., Holt I., Karamycheva S., Liang F., Quackenbush J.,
RA Keefe J.W.;
RT "Sequence evaluation of four pooled-tissue normalized bovine cDNA
RT libraries and construction of a gene index for cattle.";
RL Genome Res. 11:626-630 (2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Pooled;
RA Hartay G.P., Sonstegard T.S., Van Tassel C.P., Clawson M.L.,
RA Heaton M.P., Keefe J.W., Snelling W.M., Weidmann R.T., Smith T.P.L.;
RT "Sequencing and analysis of Bos taurus full-length insert cDNA
RT clones.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBSJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; BT021839; AAX46686.1; -; mRNA.
DR SMR; Q58CV8; 11-262.
DR Ensembl; ENSBTAG00000016032; Bos taurus.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1; 1.
DR Transferase.
KW SEQUENCE 298 AA; 34914 MW; BE97E9EA1032E37E CRC64;

Query Match 30.3%; Score 463; DB 2; Length 298;
Best Local Similarity 35.3%; Pred. No. 3.1e-30;
Matches 98; Conservative 58; Mismatches 102; Indels 20; Gaps 5;

QY 24 GVRLLPFCRCGRKMEIEANFVRPSDVMTVTPKSGTSLLEQVYLVYVLSQADPDDEIGLMNI 82
DB 19 GILQPKPTC-DTWDDQWSFQARPDLLISTYPKAGTTWTQETVDLIQGGDVNQSORAPT 77

QY 83 DEQLPVLVEYPPQ-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDL 138
DB 78 HERFFPIEWIIITPSLGSGLQANAMASPRMLKTHLPFHLLPPSFLEBNCMIYVAARNPKDN 137

QY 139 VSYVYQFHRSLRTMSYRGTFQFCRCRPMNDKLGYSWFHFVQEFWEHRMDSNVLFKLYED 198
DB 138 VSYTHFHRMEKAPPEPGTWDSEFLKFMAGEVSYNSWYHQVQEWQAKDQHRILYFYED 197

QY 199 MHRDLVTMVEQLARFLGVSCDKAQLBALTEHCQLV---DQCCNAEALP----- 244
DB 198 MKENPKHEIQKLAEPFGKSLDDKVLDTKIVDHTSFSVMKQNPMTNTSIPNMYMNLISPF 257
```

```
QY 245 VGRGRVGLWKDITVSMNEKFDLVYKQKMGKCDLTFDF 282
DB 258 MRKGVIGDWKNHFTVAQNERFDDYRKNMADTTLLHF 295

RESULT 23
STBL1_CHICK
ID STBL1_CHICK STANDARD; PRT; 296 AA.
AC Q8JG30;
DT 04-JAN-2005, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-2002, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-)
DE (Sulfotransferase 1B).
GN Name=SULT1B1; Synonyms=SULT1B;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND FUNCTION.
RC TISSUE=Lymphoid;
RX PubMed=15234270; DOI=10.1016/j.abb.2004.05.008;
RA Wilson L.A., Reynolds G.E., Darras V.M., Coughtrie M.W.H.;
RT "cDNA cloning, functional expression, and characterization of chicken
sulfotransferases belonging to the SULT1B and SULT1C families.";
RL Arch. Biochem. Biophys. 428:64-72(2004).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of many hormones,
neurotransmitters, drugs and xenobiotic compounds. Sulfonation
increases the water solubility of most compounds, and therefore
their renal excretion, but it can also result in bioactivation to
form active metabolites. Sulfates thyroid hormones including 3,3'-
diiodothyronine, catechols such as 4-nitrocatechol and 4-
isopropylcatechol and phenols such as 2-bromophenol.
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC Copyright by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AJ494980; CAD41949.1; -; mRNA.
DR HSP; P49888; IY3.
DR Ensembl; ENSGALG00000011812; Gallus gallus.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Lipid metabolism; Steroid metabolism; Transferase.
FT CHAIN
FT 1 296 Sulfotransferase family cytosolic 1B
FT member 1.
FT /FTID=PRO_0000085164.
FT NP_BIND 48 53 PAPS (By similarity).
FT FT NP_BIND 131 139 PAPS (By similarity).
FT FT NP_BIND 194 230 PAPS (By similarity).
FT FT ACT_SITE 109 109 Proton acceptor (By similarity).
FT FT BINDING 259 259 PAPS (By similarity).
SQ SEQUENCE 296 AA; 34074 MW; 487AB790416995F5 CRC64;

Query Match 30.0%; Score 459.5; DB 1; Length 296;
Best Local Similarity 35.8%; Pred. No. 6.1e-30;
Matches 96; Conservative 50; Mismatches 95; Indels 27; Gaps 4;

QY 36 EBIAFPVRPSDWIVTPKSGTSLQEVYVLVSQADPDDEITGLMNIDEQLPVLRYPOP- 94
DB 29 ERVDNFQSRPDDIVATPKSGTTSWSEIVDMILQGGDPKKCKRDAIVNRVPMLEFAAPG 88
QY 95 ----GLDIKELTSRLKSHLPYRPLSDLHNGDSKVITYMARNPKDLVSVYYQFHSRLR 150
DB 89 QMPAGTEQLENMPSRIIKTHIPADILPKSFWDKCKMIVYGRNAKVAVSYTHFDLMNK 148
QY 151 TMSYGTGTQECRRRMDKLVGWSFHVQEFWEHRMDSNVLFLKVEDMHRDLVTMVSQL 210
DB 149 LHPHPGTWQDYLEAFMAGKVAYGWSFDHVRGYWERRQBHPILYLPEDMKEDLRREKAVK 208

QY 211 ARFLGVSCDKAQLALTEHC-----HQLVQCCNAEALP-VGRGRVGL 252
DB 209 AQFLGRELTEVALDAIAHHTSPAMRDNPNSTNYSVVPVSLMDHGIS----PPMRKGIIGD 264

QY 253 WKDIETVSMNEKFDLVYKQKMGKCDLTF 280
DB 265 WKNHFTVAQSAHFDDYQYAKMGAGTDLRF 292

RESULT 24
STBL1_RAT
ID STBL1_RAT STANDARD; PRT; 291 AA.
AC P17988;
DT 01-NOV-1990, integrated into UniProtKB/Swiss-Prot.
DT 01-NOV-1990, sequence version 1.
DT 07-FEB-2006, entry version 55.
DE Sulfotransferase 1A1 (EC 2.8.2.1) (Aryl sulfotransferase) (Phenol
sulfotransferase) (PST-1) (Sulfokinase) (Aryl sulfotransferase IV)
DE (ASTIV) (Tyrosine-ester sulfotransferase) (Minoxidil
sulfotransferase).
GN Name=Sult1a1; Synonyms=Sst1a1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Murioidea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=90326537; PubMed=2374726;
RA Ozawa S., Nagata K., Gong D., Yamazoe Y., Kato R.;
RT "Nucleotide sequence of a full-length cDNA (PST-1) for aryl
sulfotransferase from rat liver.";
RL Nucleic Acids Res. 18:4001-4001(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=94131305; PubMed=8299966; DOI=10.1016/0378-1119(93)90028-2;
RA Khan A.S., Taylor B.R., Chung K., Etheredge J., Gonzales R.,
Ringer D.P.;
RT "Genomic structure of rat liver aryl sulfotransferase IV-encoding
gene.";
RL Gene 137:321-326(1993).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 7-291.
RX MEDLINE=92379794; PubMed=1511441;
RA Yerokun T., Etheredge J.L., Norton T.R., Carter H.A., Chung K.H.,
Birckbichler P.J., Ringer D.P.;
RT "Characterization of a complementary DNA for rat liver aryl
sulfotransferase IV and use in evaluating the hepatic gene transcript
levels of rats at various stages of 2-acetylaminofluorene-induced
hepatocarcinogenesis.";
RL Cancer Res. 52:4779-4786(1992).
RN [4]
RP CHARACTERIZATION.
RC TISSUE=Liver;
RX MEDLINE=93191709; PubMed=8447833;
RA Cruickshank D., Sansom L.N., Veronese M.E., Mojarrahi B.,
McManus M.E., Zhu X.;
RT "cDNA expression studies of rat liver aryl sulphotransferase.";
RL Biochem. Biophys. Res. Commun. 191:295-301(1993).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=95196680; PubMed=7889867;
RA Yamazoe Y., Ozawa S., Nagata K., Gong D.-W., Kato R.;
RT "Characterization and expression of hepatic sulfotransferase involved
in the metabolism of N-substituted aryl compounds.";
RL Environ. Health Perspect. 102:99-103(1994).
RN [6]
RP PROTEIN SEQUENCE OF 63-68, AND CHARACTERIZATION.
RX MEDLINE=95074030; PubMed=7982943;
RA Zheng Y., Bergold A., Duffel M.W.;
```





```

QY 243 LP-VGRGRVGLWKDIFTYSMNEKFDLVYKQMGKCDLTF 280
DB 254 -PFMRKGMAGDWKTTFTVAQNERFDADYAEKMGCSLSF 291

RESULT 27
Q6Y0X5_PIG
ID Q6Y0X5_PIG PRELIMINARY; PRT; 295 AA.
AC Q6Y0X5;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, sequence version 1.
DE Phenol sulfotransferase.
OS Sus scrofa (Fig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX PubMed=15014971; DOI=10.1007/s00335-002-2318-4;
RA Lin Z., Lou Y., Squires J.E.;
RT "Molecular cloning and functional analysis of porcine SULT1A1 gene and
RT its variant: a single mutation SULT1A1 causes a significant decrease
RT in sulfation activity.";
RL Mamm. Genome 15:218-226 (2004).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY193893; AAP03092.1; -; mRNA.
DR HSSP; P49891; 1AQU.
DR SMR; Q6Y0X5; 9-235.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 295 AA; 34411 MW; 1721E635F6B94EE0 CRC64;

Query Match 29.8%; Score 456; DB 2; Length 295;
Best Local Similarity 35.6%; Pred. No. 1.2e-29;
Matches 98; Conservative 51; Mismatches 108; Indels 18; Gaps 3;

QY 24 GVRLPFCRCGMEIEANFPVRPSDVWIVTPKSGTSLQEVVYLVSGQADPDEIGLMNID 83
DB 17 GVPLIKYFAEALGPLESQAMPDDVLISTPKSGTTWSEILDLIYQGDLLQKQCAPIF 76

QY 84 BOLPVLVEYQPQ-----GLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVIMYARNPKDLV 139
DB 77 VRVPLEPKIPRCPTGFELKDDTPAPRLKTHLPQLTLLPQTLDDQKVYVYARNAKDVA 136

QY 140 VSYIQPHSLRTMSYRGTFQECRFPMNDKLGYSWFEHVFQEFWEHRMDSNVFLFKYEDM 199
DB 137 VSYHYFIRMAKVYPNPGTWDSLEDFMAGEVSYSQWYQHVQGWSEIRTHPLVLYLFYEDM 196

QY 200 HRDLVTWVEQLARFLGVSQDKAQLAEALTEHC---HQLVDQCCNAEALP-----V 245
DB 197 KENPKRETKILEFVGRSLPBEETVEDIVQHTSFQEMKNAMNTYRTLSDLLDHSISAFM 256

QY 246 GRGRVGLWKDIFTYSMNEKFDLVYKQMGKCDLTF 280
DB 257 RKGITGDWKSFTVAQNERFEREADYAEKMGACNLRF 291

RESULT 28
STK2_RAT
ID STK2_RAT
AC Q9UWU9;
DT 01-DEC-2000, integrated into UniProtKB/Swiss-Prot.

```

```

DT 01-NOV-1999, sequence version 1.
DT 07-FEB-2006, entry version 29.
DE Sulfotransferase K2 (EC 2.8.2.-) (rSULT1C2A).
GN Name=Sultk2; Synonyms=Sultic2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RT TISSUE=Kidney;
RX MEDLINE=20329235; PubMed=10872834; DOI=10.1006/bbrc.2000.2744;
RA Xiangrong L., Joehnk C., Hartmann D., Schestag F., Kroemer W.,
RA Gieselmann V.;
RT "Enzymatic properties, tissue-specific expression, and lysosomal
RT location of two highly homologous rat SULT1C2 sulfotransferases.";
RL Biochem. Biophys. Res. Commun. 272:242-250(2000).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of many drugs,
CC xenobiotic compounds, hormones, and neurotransmitters.
CC -!- SUBCELLULAR LOCATION: Lysosome.
CC -!- TISSUE SPECIFICITY: Highly expressed in kidney and at lower levels
CC in stomach and liver. More specifically found in the epithelia of
CC proximal tubules of the kidney, of the bile duct, of the gastric
CC mucosa, and in hepatocytes.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AJ238392; CAB41461.1; -; mRNA.
DR HSSP; P50224; 1CJM.
DR SMR; Q9UWU9; 12-296.
DR Ensembl; ENSRNOG0000031833; Rattus norvegicus.
DR RGD; 621064; Sultic2.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Lysosome; Transferase.
FT CHAIN 1..296
FT Sulfotransferase K2.
FT NP_BIND 49..54 /FTID=PRO_0000085136.
FT NP_BIND 131..139 PAPS (By similarity).
FT NP_BIND 194..230 PAPS (By similarity).
FT NP_BIND 258..260 PAPS (By similarity).
FT ACT_SITE 109..109 Proton acceptor (By similarity).
SQ SEQUENCE 296 AA; 34840 MW; A6D0C9257C10E34B CRC64;

Query Match 29.8%; Score 456; DB 1; Length 296;
Best Local Similarity 34.7%; Pred. No. 1.2e-29;
Matches 99; Conservative 56; Mismatches 108; Indels 22; Gaps 5;

QY 16 ESKYFEFHGVRLPFCRCGMEIEANFPVRPSDVWIVTPKSGTSLQEVVYLVSGADPD 75
DB 10 QTKLKEVAGIPLRDSVDNWSQIQTFKAKPDDLICTYPKSGTTWVQIENVNIEQNGDVE 69

QY 76 EIGLWNIDQQLPVLEY---PQP-GLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVLYM 131
DB 70 KCQRTIIQHRHFFIEWARPPQPSGVVDKANAPAPRILTHLPQLLPSPFTWNNCKLYLV 129

QY 132 ARNPKDLVVSYYQFHSRLRTMSYRGTFQECRFPMNDKLGYSWFEHVFQEFWEHRMDSNV 191
DB 130 ARNAKDCWVSFYHFVRCQVLPNPGTWNEYPFTFNGKVSQSGSFHDVKGWWEIRDRYQI 189

QY 192 LFLKYEDMHRDLVTWVEQLARFLGVS-----CDKAQLEALTEHCQLVDQCCNAEALP-- 244
DB 190 LFLFYEDMKRDPKREIQKVMQFMGNLDEEVVDKIVLETSPFKMKD--NPLTFNFTIPRT 247

QY 245 -----VGRGRVGLWKDIFTYSMNEKFDLVYKQMGKCDLTF 280
DB 248 IMDQISPFMRKGIQVDMKNHFTVAQNERFEREADYAEKMGDGTSLNF 292

```



```
RESULT 29
Q95JC6 TRIVU PRELIMINARY; PRT; 296 AA.
AC Q95JC6;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE Sulfotransferase SULT1B.
OS Trichosurus vulpecula (Brush-tailed possum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.
OX NCBI_TaxID=9337;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RA Bolton-Grob R.M., McManus M.E.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBSJ databases.
CC -----
CC Copyright by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY044181; AAK72404.1; -; mRNA.
DR HSP; P50224; 1CJM.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 296 AA; 35057 MW; 4DA0BADCA70E3224 CRC64;

Query Match 29.6%; Score 453.5; DB 2; Length 296;
Best Local Similarity 35.2%; Pred. No. 1.9e-29;
Matches 94; Conservative 56; Mismatches 92; Indels 25; Gaps 3;

Qy 36 EEIANFPVRPSDWMIVTPKSTSLQEWYLVSGADPDEGLNMNIDEQLPVLEYPPQG 95
Db 29 ERIEKRWRPDIVATYPKSGTTWISVDVMDIQNDGIEKSRDAITEKPMLEFLVPG 88
Qy 96 LDI-----IKELTSRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVWSYQPHRSILR 150
Db 89 LRITGTEALEKMPSPRLIKTHLPVDLIPKDCWKNCKIYARNAKDAVSYTHDLNKK 148
Qy 151 TMSVGTQFQECRRPMNDKLGYSWFHFQEFWEHRMDSNVLFKYEDHMRDLVTMVSQL 210
Db 149 LEPYPNSWAEYLEKYVTGKVSYGSWFAHVKSMDKKTYPMLYLFYEDMKGNPKKEIEKV 208
Qy 211 ARFLGVSCDKAQLEALTEH-----CHQLVDDCCNAEALPVGRGVRGLW 253
Db 209 MQFLGKNLDEEVLEKILHHTSFEMMKNNPLNFTNISKKMDHEVSSH---LRKGIVGDW 265
Qy 254 KDIFTVMNEKFDLVYKQMGKCDLTF 280
Db 266 KNYFTVAQNEIFDAIYKEMARTILKP 292

RESULT 30
ST1A1 HUMAN
ID ST1A1 HUMAN STANDARD; PRT; 295 AA.
AC P50225; Q96U58; Q92818; Q9BVU6; Q9UGG7.
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 01-DEC-2000, sequence version 2.
DT 21-FEB-2006, entry version 59.
DE Sulfotransferase 1A1 (EC 2.8.2.1) (Aryl sulfotransferase 1) (Phenol
DE Sulfotransferase 1) (Phenol-sulfating phenol sulfotransferase 1) (P-
DE PST 1) (Thermostable phenol sulfotransferase) (Ts-PST) (HAST1/HAST2)
DE (ST1A3).
GN Name=ST1A1; Synonyms=STP, STP1; ORFNames=OK/SW-cl.88;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
```

```
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Liver;
RA Zhu X., Veronese M.E., Sansom L.N., McManus M.E.;
RL "Molecular characterisation of a human aryl sulfotransferase cDNA.";
RT Biochem. Biophys. Res. Commun. 192:671-676(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Brain;
RA Zhu X., Veronese M.E., Bernard C.C., Sansom L.N., McManus M.E.;
RL "Identification of two human brain aryl sulfotransferase cDNAs.";
RT Biochem. Biophys. Res. Commun. 195:120-127(1993).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Hippocampus;
RA Hwang S.-R., Kohn A.B., Hook V.Y.H.;
RL "Molecular cloning of an isoform of phenol sulfotransferase from human brain hippocampus.";
RT Biochem. Biophys. Res. Commun. 207:701-707(1995).
RN [4]
RP NUCLEOTIDE SEQUENCE [MRNA], AND PARTIAL PROTEIN SEQUENCE.
RC TISSUE=Liver;
RA MEDLINE=931140712; PubMed=8423770;
RL Wilborn T.W., Comer K.A., Dooley T.P., Reardon I.M., Heinrikson R.L., Falany C.N.;
RT "Sequence analysis and expression of the cDNA for the phenol-sulfating form of human liver phenol sulfotransferase.";
RL Mol. Pharmacol. 43:70-77(1993).
RN [5]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Blood;
RA MEDLINE=95209704; PubMed=7695643;
RL Jones A.L., Hagen M., Coughtrie M.W.H., Roberts R.C., Glatt H.;
RT "Human platelet phenol sulfotransferases: cDNA cloning, stable expression in V79 cells and identification of a novel allelic variant of the phenol-sulfating form.";
RL Biochem. Biophys. Res. Commun. 208:855-862(1995).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RA MEDLINE=96065417; PubMed=7581483;
RL Ozawa S., Nagata K., Shimada M., Ueda M., Tsuzuki T., Yamazoe Y., Kato R.;
RT "Primary structures and properties of two related forms of aryl sulfotransferases in human liver.";
RL Pharmacogenetics 5:S135-S140(1995).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=94306556; PubMed=8033246; DOI=10.1016/0009-2797(94)90057-4;
RL Yamazoe Y., Nagata K., Ozawa S., Kato R.;
RT "Structural similarity and diversity of sulfotransferases.";
RL Chem. Biol. Interact. 92:107-117(1994).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=96211162; PubMed=8924211;
RL Bernier F., Soucy P., Luu-The V.;
RT "Human phenol sulfotransferase gene contains two alternative promoters: structure and expression of the gene.";
RL DNA Cell Biol. 15:367-375(1996).
RN [9]
RP NUCLEOTIDE SEQUENCE.
RA PubMed=8912648; DOI=10.1006/dbrc.1996.1628;
RL Dooley T.P., Huang Z.;
RT "Genomic organization and DNA sequences of two human phenol sulfotransferase genes (STP1 and STP2) on the short arm of chromosome 16.";
RL Biochem. Biophys. Res. Commun. 228:134-140(1996).
RN [10]
RP NUCLEOTIDE SEQUENCE.
RL Rattogianis R.B., Her C., Weinshilboum R.M.;
```

Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
[11]  
RN NUCLEOTIDE SEQUENCE.  
RC TISSUE=Liver;  
RA MEDLINE=99074339; PubMed=9855620;  
RX Dajani R., Hood A.M., Coughtrie M.W.;  
RT "A single amino acid, Glu146, governs the substrate specificity of a  
RL human dopamine sulfotransferase, SULT1A3.";  
RM Mol. Pharmacol. 54:942-948 (1998).  
[12]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], AND VARIANT HIS-213.  
RC TISSUE=Colon adenocarcinoma;  
RA Shichijo S., Itoh K.;  
RT "Identification of immuno-peptidomimics that are recognized by tumor-  
RL reactive CTL generated from TIL of colon cancer patients.";  
RM Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
[13]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], AND VARIANTS HIS-213 AND  
RP VAL-223.  
RA Kainane N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,  
RX Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,  
RA Phelan M., Farmer A.;  
RT "Cloning of human full-length CDSs in BD Creator(TM) system donor  
RL vector.";  
RM Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
[14]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC TISSUE=Placenta;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RX Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Hopkins R.F., Jordan B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,  
RX Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Richards S., Worley K.C., Hale S.G., Garcia A.M., Gay L.J., Hulyk S.W.,  
RX Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
RX Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Greenwood J., Schmutz J., Myers R.M.,  
RX Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RL and mouse cDNA sequences.";  
RM Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
[15]  
RN NUCLEOTIDE SEQUENCE OF 1-107.  
RX MEDLINE=94117022; PubMed=8288252;  
RA Dooley T.P., Obermoller R.D., Leiter E.H., Chapman H.D., Falany C.N.,  
RX Deog Z., Siciliano M.J.;  
RT "Mapping of the phenol sulfotransferase gene (STP) to human chromosome  
RL 16p12.1-pl1.2 and to mouse chromosome 7.";  
RM Genomics 18:440-443 (1993).  
[16]  
RN CHARACTERIZATION.  
RX MEDLINE=94379981; PubMed=8093002;  
RA Veronesi M.E., Burgess W., Zhu X., McManus M.E.;  
RT "Functional characterization of two human sulfotransferase cDNAs that  
RL encode monoamine- and phenol-sulphating forms of phenol  
RT sulfotransferase: substrate kinetics, thermal-stability and  
RL inhibitor-sensitivity studies.";  
RM Biochem. J. 302:497-502 (1994).  
[17]  
RN MUTAGENESIS OF CYS-70.  
RX MEDLINE=94306582; PubMed=8033270; DOI=10.1016/0009-2797(94)90053-1;  
RA Falany C.N., Zhuang W., Falany J.L.;  
RT "Characterization of expressed human phenol-sulfating phenol  
RT sulfotransferase: effect of mutating cys70 on activity and  
RT thermostability.";

Chem. Biol. Interact. 92:57-66 (1994).  
[18]  
RN VARIANTS GLN-37; HIS-213 AND VAL-223.  
RX MEDLINE=98005125; PubMed=9345314; DOI=10.1006/bbrc.1997.7466;  
RA Raftogiannis R.B., Wood T.C., Otterness D.M., Van Loon J.A.,  
RX Weinshilboum R.M.;  
RT "Phenol sulfotransferase pharmacogenetics in humans: association of  
RL common SULT1A1 alleles with RS PST phenotype.";  
RM Biochem. Biophys. Res. Commun. 239:298-304 (1997).  
[19]  
RN VARIANT HIS-213.  
RX MEDLINE=20222641; PubMed=10762004;  
RX DOI=10.1097/00008571-200003000-00008;  
RA Engelte C.E., Meinl W., Boeing H., Glatt H.;  
RT "Association between functional genetic polymorphisms of human  
RL sulfotransferases 1A1 and 1A2.";  
RM Pharmacogenetics 10:163-169 (2000).  
[20]  
RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) IN COMPLEX WITH  
RP ADENOSINE-3',5'-DIPHOSPHATE (PAP) AND A PHENOL.  
RX PubMed=12471039; DOI=10.1074/jbc.M207246200;  
RA Gamage N.U., Dugleby R.G., Barnett A.C., Tresillian M., Latham C.F.,  
RX Liyou N.E., McManus M.E., Martin J.L.;  
RT "Structure of a human carcinogen-converting enzyme, SULT1A1.  
RL J. Biol. Chem. 278:7655-7662 (2003).  
CC -!- FUNCTION: Catalyzes the sulfate conjugation of catecholamines,  
CC phenolic drugs and neurotransmitters. Is also responsible for the  
CC sulfation and activation of minoxidil. Mediates the metabolic  
CC activation of carcinogenic N-hydroxyarylamines to DNA binding  
CC products and could so participate as modulating factor of cancer  
CC risk.  
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + a phenol =  
CC adenosine 3',5'-diphosphate + an aryl sulfate.  
CC -!- SUBUNIT: Homodimer.  
CC -!- SUBCELLULAR LOCATION: Cytoplasm.  
CC -!- TISSUE SPECIFICITY: Liver, lung, adrenal, brain, platelets and  
CC skin.  
CC -!- SIMILARITY: Belongs to the sulfotransferase family.  
Query Match 29.5%; Score 451; DB 1; Length 295;  
Best Local Similarity 33.8%; Pred. No. 3.1e-29;  
Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;  
QY 24 GVRPPPCRGKMEIANPVPSPDWIVTYPKSGTSLLOEVVYLVSGADPDEIGLMNID 83  
DB 17 GVPLIKYFAEALGPLQSPQAPDPLLSTYPSKSGTTWVSQILDMIYQGGDEKCHRAPIF 76  
QY 84 EOLPVLEYQPQ---GLDIIKELTSPRIKSHLPYRFLPSPDLHNGDSKVIYMARPKDLV 139  
DB 77 MRVPFLEKAPGIPSGMETLAKDTAPRLKLTHTPLALLPQTLLDQKVVYVARNAKDVA 136  
QY 140 VSYTQFHRSLRMTSYRGTFQFCRRFMDKLGYSWFEHVFWEHHRMDSNVLFUKYEDM 199  
DB 137 VSYTHFYHMAKVHPBPCTWDSFLEKFMVGEVSYGSWYQHVQEWELSRTHFPVLYFYEDM 196  
QY 200 HRDLVTWVEQLARFLGVSCDAQALEALTEHC---HQLVQCCNAELP-----V 245  
DB 197 KENPKREIQKILEFVGRSLPDEETDFVMVQHTSFKEMKKNPMNTNTYTVQPEFMDHSISPFM 256  
QY 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280  
DB 257 RKGMGADWKTFTTVAQNERFDADYAEKMGCSLSF 291  
RESULT 31  
Q2NL71 HUMAN  
ID Q2NL71 HUMAN PRELIMINARY; PRT; 295 AA.  
AC Q2NL71  
DI 07-FEB-2006, integrated into UniProtKB/TrEMBL.  
DT 07-FEB-2006, sequence version 1.  
DT 07-FEB-2006, entry version 1.  
DE SULT1A1 protein.

```
GN Name=SULT1A1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Colon;
RG NIH MGC Project;
RL Submitted (DEC-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; BC110887; AA110888.1; -; mRNA.
SQ SEQUENCE 295 AA; 34197 MW; 60E9D5BBA9159176 CRC64;

Query Match 29.5%; Score 451; DB 2; Length 295;
Best Local Similarity 33.8%; Pred. No. 3.1e-29;
Matches 93; Conservative 57; Mismatches 107; Indels 18; Gaps 3;

Qy 24 GVRLLPPFCRGMKEETANFVRPSDVWIVTPKSGTSLQEVVYLSQGADPDDELGLMID 83
Db 17 GVPLIKYFAEALGPQSFQARPDLLISTYPSKGTWVSQILDMLYQGGDLKCHRAPIF 76

Qy 84 EQLPVLEYPOP----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLV 139
Db 77 MRVPLEFKAPGIPSGMETLKDTPAPRLKTLKTHLPALLPQLTLLDQKVKVYVARNAKDVA 136

Qy 140 VSYQFHRSLRTMSYRGTFQFCRFRMNDKLGYSNFWHFHWRMDSNVFLFKYEDM 199
Db 137 VSYTHFYHMAKVHPFGTWSFLKFMVGEVSYGSYQVQVQWELSRTHPVLVLYFYEDM 196

Qy 200 HRDLVTWVEQLARFLGVCSDKAQLEALTEHC---HQLVDOCCNAEALP-----V 245
Db 197 KENPKREIQLKLEFVGRSLPDEETVDFMVQHTSFKEKMKKNPMTNTYTVQEFMDHSISPFM 256

Qy 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 257 RKGWAGDWKTTFTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 32
Q32C30 BOVIN
ID Q32C30 BOVIN PRELIMINARY; PRT; 295 AA.
AC Q32C30
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
RP NUCLEOTIDE SEQUENCE [RNA].
```

```
DT 07-MAR-2006, entry version 5.
DE Sulfotransferase, estrogen-preferring.
GN Name=STE;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Hereford; TISSUE=Reticulum;
RA Moore S., Alexander L., Brownstein M., Guan L., Lobo S., Meng Y.,
RA Tanaguchi M., Wang Z., Yu J., Prange C., Schreiber K., Shenmen C.,
RA Wagner L., Bala M., Barbazuk S., Barber S., Babakiaff R., Beland J.,
RA Chun E., Del Rio L., Gibson S., Hanson R., Kirkpatrick R., Liu J.,
RA Matsuo C., Mayo M., Santos R.R., Stott J., Tsai M., Wong D.,
RA Siddiqui A., Holt R., Jones S.J., Marra M.A.;
RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; BC102939; AA102940.1; -; mRNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 295 AA; 34660 MW; 81456185E09F016 CRC64;

Query Match 29.5%; Score 451; DB 2; Length 295;
Best Local Similarity 34.6%; Pred. No. 3.1e-29;
Matches 101; Conservative 57; Mismatches 110; Indels 24; Gaps 5;

Qy 10 STPGEFESKYFEFGVRLPPEFCRGMEIAN---PPVPSDVWIVTPKSGTSLQEVVY 66
Db 3 SKSPSFSDFYQKLGII---PMYKPFIEQHNVEFEARPDLLVIVTPKSGTSLQEVVY 59

Qy 67 LVSGQADPDDELGLMIDQLPVL----EYPOPGDIIKELTSPRLIKSHLPYRFLPSDLH 122
Db 60 MYNNGDVEKCKEDVFNRPVPLECSTEHVNGVKQLNEMASPRIVKSHLPVSLPVSFW 119

Qy 123 NGDSKVIYMARNPKDLVVSYYQFHRSLRTMSYRGTFQFCRFRMNDKLGYSNFWHFHWR 182
Db 120 EKNCKIILYLSNAKDVVSVYFFILMTVAIPDPDSFQDFVEKFMDEGVYGSWFERTKSW 179

Qy 183 WEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVCSDKAQLEALTEHC----- 231
Db 180 WEKSKNPQVLFYFEDMKENIRKVMKLEFLGRKASDELVDKIIKHTSFQEMKNPSTN 239

Qy 232 --QLVDOCCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 240 YTTLPDEVNMQKVSFFMRKGDVGNKQHTVALNEKFDHMYEQQMGKSTLKF 291

RESULT 33
ST1A1 CANFA
ID ST1A1 CANFA STANDARD; PRT; 295 AA.
AC Q29476;
DT 26-APR-2005, integrated into UniProtKB/Swiss-Prot.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 29.
DE Sulfotransferase 1A1 (EC 2.8.2.1) (Aryl sulfotransferase) (Phenol
DE sulfotransferase) (Phenol-sulfating phenol sulfotransferase) (P-PST).
GN Name=SULT1A1;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
OC Canis.
OX NCBI_TaxID=9615;
RN [1]
RP NUCLEOTIDE SEQUENCE [RNA].
```



OC Homo. NCBI\_TaxID=9606;  
 RN (1)  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Fetal liver, and Fetal spleen;  
 RX MEDLINE=97312707; PubMed=9169148; DOI=10.1006/geno.1997.4683;  
 RA Her C., Kaur G.P., Achwal R.S., Weinshilboum R.M.;  
 RT "Human sulfoltransferase SULT1C1: cDNA cloning, tissue-specific  
 expression, and chromosomal localization.";  
 RL Genomics 41:467-470(1997).  
 RN (2)  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Fetal liver;  
 RX MEDLINE=98297811; PubMed=9635888; DOI=10.1093/carcin/19.5.951;  
 RA Yoshinari K., Nagata K., Shimada M., Yamazoe Y.;  
 RT "Molecular characterization of SULT1C1-related human sulfoltransferase.";  
 RL Carcinogenesis 19:951-953(1998).  
 RN (3)  
 RP NUCLEOTIDE SEQUENCE, AND CHARACTERIZATION.  
 RC TISSUE=Fetal lung;  
 RX MEDLINE=99069375; PubMed=9852044; DOI=10.1074/jbc.273.51.33929;  
 RA Sakakibara Y., Yanagisawa K., Katafuchi J., Ringer D.P., Takami Y.,  
 RT Nakayama T., Suiko M., Liu M.-C.;  
 RL "Molecular cloning, expression, and characterization of novel human  
 SULT1C1 sulfoltransferases that catalyze the sulfonation of N-hydroxy-2-  
 acetylaminofluorene.";  
 RL J. Biol. Chem. 273:33929-33935(1998).  
 RN (4)  
 RP NUCLEOTIDE SEQUENCE, AND CHARACTERIZATION.  
 RC TISSUE=Stomach;  
 RX MEDLINE=99410886; PubMed=10481272; DOI=10.1016/S1357-2725(99)00038-2;  
 RA Hehonah N., Zhu X., Brix L., Bolton-Grob R., Barnett A., Windmill K.,  
 RT McManus M.;  
 RL "Molecular cloning, expression, localisation and functional  
 characterisation of a rabbit SULT1C2 sulfoltransferase.";  
 RL Int. J. Biochem. Cell Biol. 31:869-882(1999).  
 RN (5)  
 RP NUCLEOTIDE SEQUENCE, AND ALTERNATIVE SPLICING.  
 RX MEDLINE=20247255; PubMed=10783263; DOI=10.1006/geno.2000.6150;  
 RA Freimuth R.R., Raftogiannis R.B., Wood T.C., Moon E., Kim U.-J., Xu J.,  
 RA Siciliano M.J., Weinshilboum R.M.;  
 RT "Human sulfoltransferases SULT1C1 and SULT1C2: cDNA characterization,  
 gene cloning, and chromosomal localization.";  
 RL Genomics 65:157-165(2000).  
 RN (6)  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM SHORT).  
 RA Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,  
 RA Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,  
 RA Phelan M., Farmer A.;  
 RT "Cloning of human full-length cDNAs in BD Creator (TM) system donor  
 vector.";  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 RN (7)  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S.S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.N., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Whiting J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: Catalyzes the sulfate conjugation of many drugs,  
 CC xenobiotic compounds, hormones, and neurotransmitters. May be  
 CC involved in the activation of carcinogenic hydroxylamines. Shows  
 CC activity towards p-nitrophenol and N-hydroxy-2-acetylami-  
 CC fluorene (N-OH-2AAF).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasm.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=Short; Name=Long;  
 CC IsoId=000338-1; Sequence=Displayed;  
 CC IsoId=000338-2; Sequence=VSP\_006303;  
 CC -!- TISSUE SPECIFICITY: Found in adult stomach, kidney and thyroid  
 CC gland, and in fetal kidney and liver.  
 CC -!- SIMILARITY: Belongs to the sulfoltransferase family.  
 CC  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC  
 CC EMBL: U60036; AAC51285.1; -; mRNA.  
 CC EMBL: AB008164; BAA28346.1; -; mRNA.  
 CC EMBL: AF026303; AAC00409.1; -; mRNA.  
 CC EMBL: AF186251; AAF72799.1; -; mRNA.  
 CC EMBL: AF186252; AAF72800.1; -; mRNA.  
 CC EMBL: AF186253; AAF72801.1; -; mRNA.  
 CC EMBL: AF186254; AAF72802.1; -; mRNA.  
 CC EMBL: AF186255; AAF72803.1; -; mRNA.  
 CC EMBL: AF186256; AAF72804.1; -; mRNA.  
 CC EMBL: AF186257; AAF72805.1; -; Genomic DNA.  
 CC EMBL: AF186258; AAF72805.1; JOINED; Genomic DNA.  
 CC EMBL: AF186260; AAF72805.1; JOINED; Genomic DNA.  
 CC EMBL: AF186261; AAF72805.1; JOINED; Genomic DNA.  
 CC EMBL: AF186262; AAF72806.1; -; Genomic DNA.  
 CC EMBL: AF186258; AAF72806.1; JOINED; Genomic DNA.  
 CC EMBL: AF186259; AAF72806.1; JOINED; Genomic DNA.  
 CC EMBL: AF186260; AAF72806.1; JOINED; Genomic DNA.  
 CC EMBL: AF186261; AAF72806.1; JOINED; Genomic DNA.  
 CC EMBL: BT006951; AAP35597.1; -; mRNA.  
 CC EMBL: BC005353; AAH05353.1; -; mRNA.  
 CC PDB: 2ETG; X-ray; A/B=1-296.  
 CC Ensemble; ENSG00000198203; Homo sapiens.  
 CC H-InvDB; HIX0002353; -.  
 CC HGNC; HGNC:11456; SULT1C1.  
 CC MIM; 602385; gene.  
 CC GO; GO:0005737; C:cytoplasm; TAS.  
 CC GO; GO:0008146; F:sulfoltransferase activity; TAS.  
 CC GO; GO:0009308; P:amine metabolism; TAS.  
 CC InterPro; IPR000863; Sulfoltransferase.  
 CC Pfam; PF00685; Sulfoltransferase 1; 1.  
 CC ProDom; PD001218; Sulfoltransferase; 1.  
 CC 3D-structure; Alternative splicing; Polymorphism; Transferase.  
 CHAIN 1 296 Sulfoltransferase ICI.  
 FT /FTId=PRO\_0000085132.  
 FT NP\_BIND 49 54 PAPS (By similarity).  
 FT NP\_BIND 131 139 PAPS (By similarity).  
 FT NP\_BIND 194 230 PAPS (By similarity).  
 FT NP\_BIND 258 260 PAPS (By similarity).  
 FT ACT\_SITE 109 109 Proton acceptor (By similarity).  
 FT VARSPLIT 93 113 SSNPASTSQAKITD (in isoform Long).  
 FT /FTId=vsp\_006303.  
 FT Y -> H (in dbSNP:17036091).  
 FT /FTId=VAR\_021986.  
 FT S -> A (in dbSNP:17036104).  
 FT /FTId=VAR\_021987.

FT	HELIX	29	33
FT	TURN	34	34
FT	TURN	39	40
FT	STRAND	42	46
FT	TURN	48	49
FT	STRAND	50	51
FT	HELIX	52	64
FT	TURN	65	65
FT	STRAND	82	82
FT	TURN	83	84
FT	STRAND	88	89
FT	TURN	90	90
FT	HELIX	93	99
FT	STRAND	100	100
FT	STRAND	102	103
FT	STRAND	105	108
FT	TURN	112	114
FT	STRAND	115	115
FT	TURN	117	117
FT	HELIX	118	121
FT	TURN	122	123
FT	STRAND	125	130
FT	HELIX	133	146
FT	TURN	148	149
FT	STRAND	150	150
FT	STRAND	155	155
FT	HELIX	156	164
FT	TURN	165	166
FT	STRAND	168	169
FT	TURN	170	170
FT	HELIX	173	183
FT	TURN	184	186
FT	STRAND	187	193
FT	HELIX	194	199
FT	HELIX	201	211
FT	TURN	212	213
FT	HELIX	218	227
FT	TURN	228	228
FT	STRAND	229	229
FT	HELIX	230	235
FT	STRAND	262	262
FT	TURN	264	264
FT	HELIX	265	268
FT	HELIX	271	285

Query Match 29.3%, Score 449; DB 1; Length 296;

Best Local Similarity 33.1%, Pred.No.4.6e-29;

Matches 94; Conservative 61; Mismatches 103; Indels 24; Gaps 4

Qy	18	KYF	BFGH	YRLP	FFCR	GKMEEI	ANFP	VRPS	DVMIV	ITYPK	SGTSL	LQEV	YLV	VSQG	ADP	DEI	77
Db	12	KL	KEVE	GTLL	QPAT	VDNWS	QIQS	PEAK	PDLL	ICTY	PKAG	ITWIQ	IEIV	DMIE	QNGD	VEK	71
Qy	78	GL	MNID	EQLP	VLEY	---	POP	-GLD	IKEL	TSPRL	IKSHL	PYRFL	PSDL	HLNG	DSKV	IYMAR	133
Db	72	QR	AIQHR	HPPIE	WARP	PPQSG	VEKAK	AMPSP	PRILK	THLS	TQLL	PPSP	FWENN	CKFL	YVAR	131	
Qy	134	NP	KDLV	SVSY	YPHS	RLR	TMS	YRGTF	QFEC	RCR	FMND	KLGY	SGWF	EHVQ	BFWE	HMDSN	193
Db	132	NA	KCMV	SYIY	YFQR	MN	HMLP	DP	GTWEE	YF	FEI	FING	KV	WGSF	DFHVG	KWEMKDR	191
Qy	194	LK	YEDM	HRDL	VTWV	EQ	LARFI	LG	VS	CDK	KAQ	LEAL	TEH	-----	CHOL	VDQ	236
Db	192	LF	YED	IKRDP	KHEIR	KVMQ	FMGK	KYD	ETV	LDK	IVQ	ETSF	EK	MKEN	PMTNR	STVSK	251
Qy	237	CC	NAEAL	PVGR	VG	RLW	KDIF	TVS	MNEK	FDL	VYK	OKM	GKCD	LTF	280		
Db	252	SIS	SF--	--MR	KGTV	GDW	KNHT	V	QA	NERF	DEI	YRR	MEG	TS	INF	292	

RESULT 36

ST1C2 HUMAN STANDARD; PRT: 302 AA.

ID ST1C2 HUMAN

DR EMBL; AF055584; AAC95519.1; -; mRNA.  
DR EMBL; AF186263; AAF72810.1; -; Genomic DNA.  
DR EMBL; AC068941; AAY14742.1; -; Genomic DNA.  
DR PDB; 2AD1; X-ray; A=7-302.  
DR Ensembl; ENSG00000198075; Homo sapiens.  
DR HGNC; HGNC:11457; SULT1C2.  
DR MIM; 608357; gene.  
DR LinkHub; O75897; -.  
DR GO; GO:0008146; F:sulfotransferase activity; NAS.  
DR InterPro; IPR000863; Sulfotransferase.  
DR Pfam; PF00685; Sulfotransferase 1; 1.  
DR ProDom; PD001218; Sulfotransferase; 1.  
KW 3D-structure; Polymorphism; Transferase.  
FT CHAIN 1 302 Sulfotransferase 1C2.  
FT FT /FTID=PRO\_0000085137.  
FT NP\_BIND 55 60 PAPS (By similarity).  
FT NP\_BIND 137 145 PAPS (By similarity).  
FT NP\_BIND 200 236 PAPS (By similarity).  
FT NP\_BIND 264 286 PAPS (By similarity).  
FT ACT\_SITE 115 115 Proton acceptor (By similarity).  
FT VARIANT 5 5 D -> E (in dbSNP:1402467).  
FT /FTID=VAR\_025404.  
SQ SEQUENCE 302 AA; 35520 MW; 905E1820A5E222F3 CRC64;  
Query Match 29.3%; Score 449; DB 1; Length 302;  
Best Local Similarity 36.2%; Pred. No. 4.7e-29;  
Matches 96; Conservative 52; Mismatches 99; Indels 18; Gaps 3;  
QY 36 EEIANFPVRPSDVIVTPKSGTSLLOEVVYVLSQADPEIGLNMIDQLPVLVYVQP- 94  
DB 36 DKIMFQAKPDDLSTYPAKGTWTQEIIVELIQNEGVDEKSKRAPHQRFPFLEMKIPS 95  
QY 95 ---GLDIIKELSPRLIKSHLYRFLPSDLHNGSKVIYMARNPKDLVYVYQFHSRLT 151  
DB 96 LSGLEQAHAMPSRLIKTHLPFLLPPSLLEKNCKLIIVARNPKDNVSYTHFORMNKA 155  
QY 152 MYRGTQFQFCRRFNMNDKLGYSWFHEHQEFWEHRMDSNVLFKYEDMHRDLVTWVQOLA 211  
DB 156 LPAPGTWBYEFTFLAGKVCWSWHEHVKGMWEAKDKHILYLFYEDMKKNPKHEIQKLA 215  
QY 212 RFLGVSCDKAQLEALTEHQVLFV---DQCNAEALP-----VGRGVRGLWKDIF 257  
DB 216 EFIGKKLDKVDKIVHTSYFDVMQKNPMANYSSIPAEIMDSHSIPFMRKKGAVGDKKHF 275  
QY 258 TVSMNEKFDLVYKQKMGKCDLTFDF 282  
DB 276 TVAQNERFEDYKKKMTDRLTFHP 300  
RESULT 37  
Q95MF8\_PIG PRELIMINARY; PRT; 295 AA.  
ID Q95MF8\_PIG  
AC Q95MF8;  
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.  
DT 01-DEC-2001, sequence version 1.  
DT 07-FEB-2006, entry version 14.  
DE Estrogen sulfotransferase.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;  
OC Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22345564; PubMed=12457956; DOI=10.1016/S0739-7240(02)00172-8;  
RA Kim J.G., Vallet J.L., Rohrer G.A., Christenson R.K.;  
RT "Characterization of porcine estrogen sulfotransferase.";  
RL Domest. Anim. Endocrinol. 23:493-506(2002).  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; AF389855; AAK72967.1; -; mRNA.

DR HSSP; P49888; 1HY3.  
DR SMR; Q95MF8; 4-293.  
DR GO; GO:0008146; F:sulfotransferase activity; IEA.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR InterPro; IPR000863; Sulfotransferase.  
DR Pfam; PF00685; Sulfotransferase 1; 1.  
DR ProDom; PD001218; Sulfotransferase; 1.  
KW Transferase.  
SQ SEQUENCE 295 AA; 34940 MW; 7677C28A91B38167 CRC64;  
Query Match 29.2%; Score 447; DB 2; Length 295;  
Best Local Similarity 34.3%; Pred. No. 6.7e-29;  
Matches 96; Conservative 56; Mismatches 106; Indels 22; Gaps 4;  
QY 16 ESKYFEF---HGVLPPFCRGKMEIAPFPRSDVMIVTPKSGTSLLOEVVYVLSQG 71  
DB 5 KSAVLDYFGRINGILLYKPFIEYNDVETPEARDDLVIVTPKSGTGWSEIVVMYTE 64  
QY 72 ADPEIGLNMIDQLPVL-----EYQPGCLDIKELTSPRLIKSHLYRFLPSDLHNGSK 127  
DB 65 GDVEKCKEDTTFNRPYLECETENVMNGVKOLQMASPRIVKSHLPPELLEPVSEWKNCK 124  
QY 128 VIYNARNPKDLVSVYQFHSRLTMSYRGTQFCRRFMNDKLGYSWFHEHQEFWEHRM 187  
DB 125 IIVCRNAKVVVSYIYFFFLMTANPDGSPQDFVEKFMGGEVPGVSKYTKTSWWEKRT 184  
QY 188 DSNVLFKYEDMHRDLVTWVQLARFLGVSCDKAQLEALTEHCH-----QLV 234  
DB 185 NPQVLFIYEDMKENIRKEVNRLEIFLGRKASDELVDKIHKHTSFQEMKNNPSTNYTLP 244  
QY 235 DQCNAEALP-VGRGVRGLWKDIFTVSMNEKFDLVYKQKM 273  
DB 245 DEVMNQKVSAPMRKGIAGDKWKNYFTVALNEKEDIHYEQQM 284  
RESULT 38  
ST1B1\_RAT STANDARD; PRT; 299 AA.  
ID ST1B1\_RAT  
AC P52847;  
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.  
DT 15-DEC-1998, sequence version 2.  
DT 07-FEB-2006, entry version 42.  
DE Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-)  
DE (Sulfotransferase 1B) (DOPA/tyrosine sulfotransferase).  
GN Name=St1b1; Synonyms=St1b1;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridea; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [MRNA], AND PROTEIN SEQUENCE OF 108-117; 126-134  
RP AND 260-266.  
RC STRAIN=Sprague-Dawley; TISSUE=Liver;  
RX MEDLINE=96107201; PubMed=8530477; DOI=10.1074/jbc.270.51.30470;  
RA Sakakibara Y., Takami Y., Zwieb C., Nakayama T., Suiko M.,  
RA Nakajima H., Liu M.-C.;  
RT "Purification, characterization, and molecular cloning of a novel rat  
RT liver Dopa/tyrosine sulfotransferase.";  
RL J. Biol. Chem. 270:30470-30478(1995).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [MRNA].  
RX MEDLINE=94306556; PubMed=8033246; DOI=10.1016/0009-2797(94)90057-4;  
RA Yamazoe Y., Nagata K., Ozawa S., Kato R.;  
RT "Structural similarity and diversity of sulfotransferases.";  
RL Chem. Biol. Interact. 92:107-117(1994).  
RN [3]  
RP NUCLEOTIDE SEQUENCE [MRNA].  
RC TISSUE=Liver;  
RX MEDLINE=98104061; PubMed=9443824;  
RA Fujita K., Nagata K., Ozawa S., Sasano H., Yamazoe Y.;  
RT "Molecular cloning and characterization of rat ST1B1 and human ST1B2  
RT cDNAs, encoding thyroid hormone sulfotransferases.";

```

RL J. Biochem. 122:1052-1061 (1997).
RN [4]
RP FUNCTION.
RX PubMed=12773305; DOI=10.1152/ajpendo.00046.2003;
RA Kester M.H., Kaptein E., Roest T.J., van Dijk C.H., Tibboel D.,
RA Meinel W., Glatz H., Coughtrie M.W., Visser T.J.;
RT "Characterization of rat iodothyronine sulfotransferases.";
RL Am. J. Physiol. 285:E592-E598(2003).
CC -|- FUNCTION: Catalyzes the sulfate conjugation of many hormones,
CC neurotransmitters, drugs and xenobiotic compounds. Sulfonation
CC increases the water solubility of most compounds, and therefore
CC their renal excretion, but it can also result in bioactivation to
CC form active metabolites. Sulfation of DOPA, tyrosine isomers and
CC thyroid hormones such as 3,3',5-triiodothyronine and 3,3'-
CC diiodothyronine. May play a role in the limitation of the
CC production of L-DOPA and L-m-tyrosine and also in facilitating
CC their excretion.
CC -|- SUBUNIT: Monomer.
CC -|- SUBCELLULAR LOCATION: Cytoplasm.
CC -|- PTM: The N-terminus is blocked.
CC -|- SIMILARITY: Belongs to the sulfotransferase family.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U38419; AAC52387.1; -; mRNA.
DR EMBL; D89375; BAA24546.1; -; mRNA.
DR PIR; P50224; 1CJM.
DR HSSP; P50224; 1CJM.
DR SMR; P52847; 5-296.
DR Ensembl; ENSRNOG000001967; Rattus norvegicus.
DR RGD; 708534; Sulf1b1.
DR GO; GO:0005829; Cytochrome.
DR GO; GO:0004062; F:aryl sulfotransferase activity; IDA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR Direct protein sequencing; Lipid metabolism; Steroid metabolism;
KW Transferase.
FT CHAIN 1 299 Sulfotransferase family cytosolic 1b
FT member 1.
FT NP_BIND 48 53 /FTID=PRO_0000085163.
FT NP_BIND 131 139 PAPS (By similarity).
FT NP_BIND 194 230 PAPS (By similarity).
FT ACT_SITE 109 109 PAPS (By similarity).
FT BINDING 259 259 Proton acceptor (By similarity).
FT CONFLICT 68 68 E -> G (in Ref. 1).
FT SEQUENCE 299 AA; 34835 MW; FDECEC304EE788A8 CRC64;
SQ
Query Match 29.2%; Score 446.5; DB 1; Length 299;
Best Local Similarity 36.0%; Pred. No. 7.5e-29;
Matches 95; Conservative 52; Mismatches 98; Indels 19; Gaps 3;
QY 36 BEIANFPVRPSDVWVTPKSGTSLQEVVYLVSGADPDEIGLNMIDBOLPVLEYPQP- 94
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 29 EXIEFQSRPCDIVPTTPKSGTTLWSEIVDLVNDGVNVEKCRDVITSKVPMLEQNVPG 88
QY 95 -----GLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDKLVVSYQPHRSLR 150
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 89 ARRSVGLLKKTPSPRIIKTHLPIDLLPKSFDWNCCKMIYLARNKGVAVSYHFDLMNN 148
QY 151 TMSYRGTFQFCRRFMDKLGVSFEHVFQEFWEHRMDSNVLFKYEDHRLVTWVQQL 210
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 149 IQPLFGTWEEYLEKFLAGNAVYGFHDVYKSWKEKREGHPILFYLEDLKNPKKEIKKI 208
QY 211 ARFLGVSCDKAQLEALTECHQLV---DQCCNAEALP-----VGRGRVGLWKDI 256
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 209 ANFLDKTLDEHTLERIVHTSEVMDKNPLVNYTHPLTEIMDHKSPPFMRKGVVDGKNY 268
QY 257 FTVSMNEKFDLVYKQKMGKCDLTF 280
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 269 FTMTQSEKFDIYKKKLSGTTLEF 292

```

---

```

RESULT 39
STIAL_MACFA
ID STIAL_MACFA STANDARD; PRT; 295 AA.
AC P52846;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-1996, sequence version 1.
DT 07-FEB-2006, entry version 38.
DE Sulfotransferase 1A1 (SC 2.8.2.1) (Aryl sulfotransferase) (Phenol
DE sulfotransferase) (Phenol-sulfating phenol sulfotransferase) (P-PST).
GN Name=SLT1A1; Synonyms=STP;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Liver;
RA Ogura K., Satsukawa M., Okuda H., Watabe T.;
RT "cDNA cloning and bacterial expression of monkey liver phenol
RT sulfotransferase.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -|- FUNCTION: Catalyzes the sulfate conjugation of catecholamines and
CC phenolic drugs and neurotransmitters (By similarity).
CC -|- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + a phenol =
CC adenosine 3',5'-bisphosphate + an aryl sulfate.
CC -|- SUBUNIT: Homodimer (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -|- SIMILARITY: Belongs to the sulfotransferase family.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; D85514; BAA12822.1; -; mRNA.
DR PIR; G02924; G02924.
DR HSSP; P50224; 1CJM.
DR SMR; P52846; 8-295.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR Catecholamine metabolism; Lipid metabolism; Steroid metabolism;
KW Transferase.
FT CHAIN 1 295 Sulfotransferase 1A1.
FT NP_BIND 48 53 /FTID=PRO_0000085129.
FT NP_BIND 130 138 PAPS (By similarity).
FT NP_BIND 193 229 PAPS (By similarity).
FT NP_BIND 257 259 PAPS (By similarity).
FT ACT_SITE 108 108 Proton acceptor (By similarity).
FT SEQUENCE 295 AA; 34239 MW; BD51639D1570A841 CRC64;
SQ
Query Match 29.2%; Score 446; DB 1; Length 295;
Best Local Similarity 33.8%; Pred. No. 8.2e-29;
Matches 93; Conservative 56; Mismatches 108; Indels 18; Gaps 3;
QY 24 GVRLLPPFCRGKMEIANFPVRPSDVWVTPKSGTSLQEVVYLVSGADPDEIGLNMID 83
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 17 GVPLIKYFAEALGPLQSFRRPDDLLISTYKPSGTWVSQILDMYQGGDLKCRRAPIF 76
QY 84 EQLPVLEYPQP-----GLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDKLV 139
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 77 MRVFPLEPKVPGISGEMETLKDTFAPRLKTHLPALLQTLLOKQKVTVVARNAKOVA 136
QY 140 VSYTQFHRSLTMSYRGTFQFCRRFMDKLGVSFEHVFQEFWEHRMDSNVLFKYEDM 199
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 137 VSYHYHFHMAKVHPEPGTWDSFLEKFWAGEVSYGSYGVHGVWELSHTHPVLVLYFEDM 196
QY 200 HRDLVTWVEQLARFLGVSCDKAQLEALTEHC---HQLVDQCCNAEALP-----V 245
DQ |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 197 KENPKREIKLEFVGRSLPEETVDMVQHTSFKMKGNPMANYTTIPQELMDHSISPFM 256

```



FT	ACT SITE	108	108	Proton acceptor (By similarity).
FT	MOD RES	216	228	Phosphoserine (by PKA) (Potential).
FT	MOD RES	228	216	Phosphoserine (by PKA) (Potential).
FT	CONFLICT	117	117	S -> E (in Ref. 2).
FT	CONFLICT	171	171	S -> E (in Ref. 2).
FT	CONFLICT	248	248	M -> Q (in Ref. 2).
SEQ	SEQUENCE	295 AA;	34640 MW;	69147C73146913FD CRC64;

  

Query Match	29.2%;	Score 446;	DB 1;	Length 295;
Best Local Similarity	34.2%;	Pred. No. 8.2e-29;		
Matches 100;	Conservative 57;	Mismatches 111;	Indels 24;	Gaps 5;

  

QY	10	STPGEPESKYFPHGVRLPPFCRKGMEETAN----	FPVRPSDVMIVTPKSGTSLQGVVY	66
DB	3	SSKPSFSDFGKUGGI----	PMYKFIQEHNVVEFEARDDLVIVTPKSGTTLWLSIIC	59
QY	67	LVSQGAADPDEIGLNMIDELQPLV----	EYPQGLDIIKELTSRLIKSHLPYRFLPSDLH	122
DB	60	MYNNGDVEKREDVIFNRVPYLECSTEHWKMGVKQLNEMASPRIVKSHLPVKLLPVSVF	119	
QY	123	NGDSKVIYARNPKDLVSVYQPHRSILRTMSYRGTTQEFCCRRMNDKLGYGSGWFEHVQEF	182	
DB	120	EKNCKIYILSRNAKDVVSVYFYLLMVTAI	PPDPSFQDFVEKFMDEGVPYGSWFHTKSW	179
QY	183	WEHRMDSNVFLKYEDMHDRDLVTMVQOLARFLGVSCDKAQLEALTEHCH-----	231	
DB	180	WEKSKNPQVLFYEDMDENIRKENVKLLLEFLGRKASDELVDKIIKHTSFQEMKNPSTN	239	
QY	232	--QLVDCCNAEALP--VGRGRVGLWKDIFTVMNEKFDLVYKQKMGKCDLTF	280	
DB	240	YTTLPDEVNNQKVSPPMRKGDVGDWKNHFTVALNEKEDMHYEQOMKGSTLKF	291	

  

RESULT 41	ST1B1 HUMAN
ID	ST1B1 HUMAN
AC	Q43704; O15497; Q96F11; Q9UK34; PRT; 296 AA.
DT	04-JAN-2005, integrated into UniProtKB/Swiss-Prot.
DT	04-JAN-2005, sequence version 2.
DT	07-FEB-2006, entry version 31.
DE	Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-)
DE	(Sulfotransferase 1B2) (Thyroid hormone sulfotransferase).
GN	Names=ST1B1; Synonyms=ST1B2, SULT1B2;
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC	Homo
OX	NCBI_TaxID=9606;
RN	[1]
RP	NUCLEOTIDE SEQUENCE, AND FUNCTION.
RC	TISSUE=Liver;
RX	MEDLINE=98104061; PubMed=9443824;
RA	Fujita K., Nagata K., Ozawa S., Sasano H., Yamazoe Y.;
RA	"Molecular cloning and characterization of rat ST1B1 and human ST1B2
RT	cDNAs, encoding thyroid hormone sulfotransferases.";
RL	J. Biochem. 122:1052-1061(1997).
RN	[2]
RP	NUCLEOTIDE SEQUENCE, FUNCTION, AND TISSUE SPECIFICITY.
RC	TISSUE=Liver;
RX	MEDLINE=98130699; PubMed=9463486;
RA	Wang J., Falany J.L., Falany C.N.;
RT	"Expression and characterization of a novel thyroid hormone-sulfating
RT	form of cytosolic sulfotransferase from human liver.";
RL	Mol. Pharmacol. 53:274-282(1998).
RN	[3]
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC	TISSUE=Bone marrow;
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Detge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,



Db	16	HGYPMTCAPASNWEKIEQFHSRPDDDIVIATYPKSGTTWVSEIIMDLNDGDTEKCKRGFI	75
Qy	83	DEQLPVLEYPPOGL-----DIIKELTSRLIKSHLPYRFLPSDLHGSDKVTYMARNPKD	137
Db	76	TEKVPMEWTLPLGLRTSGIEQLEKNSPRIVKTHLPTDLLPKSFWENNCMIYLARNAKD	135
Qy	138	LVVSYYQHRSLRUTMSYRGTFQFCRPMNDKLGYSWFPEHVFQEFWEHRMDSNVFLKYE	197
Db	136	VVSYSYHFDLMNNLPQFGTMBEYLEKEFLTGKVAYGSMFTHVKNMKCKEEHPILFLYYE	195
Qy	198	DMHRDLVTWVEQLARFLGVSCDKAQLAALTECHQLV---DQCCNAEALP-----	244
Db	196	DMKENPKZEIKKIIFLEKNLNDLDRILIHHTSFVVMKNDPLVNYTHLPTTVMDSKSP	255
Qy	245	-VGRGRVLWKRDIFTVSMNEKFPDLVYKQKMGKCDLTF	280
Db	256	FMRKGTAGDKNYFTVAONEKFDALYETEMSKALOF	292

RESULT 42		
ST1B1	CANFA	
ID	ST1B1	CANFA STANDARD; PRT; 296 AA.
AC	Q95JD5;	
DT	04-JAN-2005,	integrated into UniProtKB/Swiss-Prot.
DT	01-DEC-2001,	sequence version 1.
DT	07-FEB-2006,	entry version 22.
DE	Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-)	
DE	(Sulfotransferase 1B) (CSULT1B1).	
GN	Name=SLUT1B1; Synonyms=ST1B2;	
OS	Canis familiaris (Dog).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;	
OC	Canis.	
OX	NCBI_TaxID=9615;	
RN	[1]	
RP	NUCLEOTIDE SEQUENCE [MRNA], FUNCTION, AND TISSUE SPECIFICITY.	
RC	TISSUE=Liver;	
RX	MEDLINE:12661348; PubMed=11368519; DOI=10.1006/abbi.2001.2373;	
RA	Tsai C., Falany C.N., Morgenstern R., Swedmark S.;	
RT	"Molecular cloning, expression, and characterization of a canine	
RT	sulfotransferase that is a human ST1B2 ortholog.";	
RL	Arch. Biochem. Biophys. 390:87-92(2001).	
CC	-!- FUNCTION: Catalyzes the sulfate conjugation of many hormones,	
CC	neurotransmitters, drugs and xenobiotic compounds. Sulfonation	
CC	increases the water solubility of most compounds, and therefore	
CC	their renal excretion, but it can also result in bioactivation to	
CC	form active metabolites. Sulfates thyroid hormones including 3,3'-	
CC	diiodothyronine, 3,5,3'-triiodothyronine and 3,5'-diiodothyronine	
CC	(lesser degree) and phenols such as 1-naphthol and p-nitrophenol.	
CC	-!- SUBUNIT: Monomer (By similarity).	
CC	-!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).	
CC	-!- TISSUE SPECIFICITY: Expressed highly in the colon, kidney and	
CC	small intestine of male and female dogs. Highly expressed in the	
CC	jejunum and ileum of the male dog than the female dog, which	
CC	displayed more expression in duodenum (at protein level).	
CC	-!- SIMILARITY: Belongs to the sulfotransferase family.	
CC	Copyrighted by the UniProt Consortium, see <a href="http://www.uniprot.org/terms">http://www.uniprot.org/terms</a>	
CC	Distributed under the Creative Commons Attribution-NoDerivs License	
CC	-----	
EMBL	AY004332; AAF86583.1; -; mRNA.	
DR	SMR; Q95JD5; 4-296.	
DR	Ensembl; ENSCAFG0000002871; Canis familiaris.	
DR	InterPro; IPR000863; Sulfotransferase.	
DR	Pfam; PF00688; Sulfotransferase; 1.	
DR	ProDom; PD001218; Sulfotransferase; 1.	
KW	Lipid metabolism; Steroid metabolism; Transference.	
FT	CHAIN 1 296 Sulfotransferase family cytosolic 1B	
FT	member 1	
FT	/FTID=PRO_0000085160.	
FT	NP_BIND 48 53 PAPS (By similarity).	
FT	NP_BIND 131 139 PAPS (By similarity).	
FT	NP_BIND 194 230 PAPS (By similarity).	
FT		

FT	ACT SITE	109	109		Proton acceptor (By similarity).
FT	BINDING	259	259		PAPS (By similarity).
SQ	SEQUENCE	296 AA;	34869 MW;	3D54CF003BF09365 CRC64;	
	Query Match	29.0%;	Score 443;	DB 1:	Length 296;
	Best Local Similarity	34.8%;	Fred. No. 1.5e-28;		
	Matches 101;	Conservative 51;	Mismatches 118;	Indels 20;	Gaps 5
QY	11	TPGFESKYF-FHGVRPPFCRGMESIANFPVPSPDVMTVITPKSGTSLQLQEIVVLVS	69	:	:
Dd	3	SPKDFLRKLAKWIHGYPIIYTFANNWENIEQHFSRPDDIIATIPKSGCTTWSEIVDMVL	62	:	:
QY	70	QGAPDDEIGLMNIDQLPVLEYPOPL-----DIIKELTSPLRIKSHLPYRFLPSDLNG	124	:	:
Dd	63	NNGDVEKCKDRFITVKVPMLEMAVPLRGTSIQEULEKNPSPRLVKTHLPIALLPKSFWEN	122	:	:
QY	125	DSKVITYWARNPGLVSVSYQPHRSIRLTMSYRGTFQEFCCRPFMNDKLGYSMPFEHVQEFWE	184	:	:
Dd	123	NCKMIYLARNAKDAVSYHYHFDLNNULPEAPGPWEYLERFMGTGNVAYGSMFNHVKSWMK	182	:	:
QY	185	HMRSDNVLFYLYEDMDHLDLVMTVBQLARFLVGSCDKAQLEALTEHC--HQLVDQ--CCNAE	241	:	:
Dd	183	KKEEHPILFLYYEDMKENPKREVQKIARFLEKLNDELVDKIHHSTFEMMKCNPLVNVT	242	:	:
QY	242	ALP-----VGRGRVLGWDIFTVSNNEKXPDLYVKOMKGCDLTF	280	:	:
Dd	243	HLPSTVMDSKSSFMKGIAGDWKNYFTVAQNEKFDVIYKKEMSGTTLQF	292	:	:
RESULT 43	STIC1_RABBIT				
ID	STIC1_RABBIT	STANDARD;	PRT;	296 AA.	
AC	046503;				
DT	01-DEC-2000,	integrated into UniProtKB/Swiss-Prot.			
DT	01-JUN-1998,	sequence version 1.			
DT	07-FEB-2006,	entry version 33.			
DE	Sulfotransferase 1C1 (EC 2.8.2.-)	(rabsULT1C2).			
GN	Name=ULT1C1;				
OS	Oryctolagus cuniculus (Rabbit).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;				
OC	Oryctolagus.				
RN	NCBI_TaxID=9986;				
OX	[1]				
RP	NUCLEOTIDE SEQUENCE [mRNA], AND CHARACTERIZATION.				
RC	TISSUE=Stomach;				
RX	MEDLINE=99410886; PubMed=10481272; DOI=10.1016/S1357-2725(99)00038-2;				
RA	Henhonah N., Zhu X., Brix L., Bolton-Grob R., Barnett A., Windmill K., McManus M.;				
RT	"Molecular cloning, expression, localisation and functional characterisation of a rabbit SULT1C2 sulfotransferase.";				
RL	Int. J. Biochem. Cell Biol. 31:869-882(1999).				
CC	-! FUNCTION: Catalyzes the sulfate conjugation of many drugs, xenobiotic compounds, hormones, and neurotransmitters. Shows activity towards p-nitrophenol. Probably the major stomach sulfotransferase.				
CC	-! SUBCELLULAR LOCATION: Cytoplasm.				
CC	-! TISSUE SPECIFICITY: Found in gastrointestinal tract tissues, liver and kidney.				
CC	-! SIMILARITY: Belongs to the sulfotransferase family.				
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms				
CC	Distributed under the Creative Commons Attribution-NoDerivs License				
CC	EMBL; AF026304; AAC00410.1; -; mRNA.				
DR	HSSP; P50224; 1CJM.				
DR	SMR; O46503; 12-296.				
DR	InterPro; IPR000863; Sulfotransferase.				
DR	Pfam; PF00685; Sulfotransferase 1; 1.				
DR	ProDom; PD001218; Sulfotransferase; 1.				
KW	Transferase.				
FT	CHAIN	1	296		Sulfotransferase 1C1. /FTId=PRO_0000085134.

FT	NP_BIND	49	54	PAPS (By similarity).
FT	NP_BIND	131	139	PAPS (By similarity).
FT	NP_BIND	194	230	PAPS (By similarity).
FT	NP_BIND	258	260	PAPS (By similarity).
FT	ACT_SITE	109	109	Proton acceptor (By similarity).
SEQ	SEQUENCE	296 AA; 34572 MW; 41089DD0F8D4F295 CRC64;		

Query Match 28.9%; Score 442.5; DB 1; Length 296;  
 Best Local Similarity 33.1%; Pred. No. 1.6e-28;  
 Matches 95; Conservative 66; Mismatches 107; Indels 19; Gaps 6

QY	12	PGCESKYFEPHGVRLLPFPFCGKKEEIANFPVRSDWIVTPYKSGTSLQGEVYLVSQG	71
DB	7	PGK-OTQLREVGEVPLQAAIVDNNQIQSFQAKPDDLICTYPKSGTGTWIOEIVDMIQN	65
QY	72	ADPDEIGLIMNIDQLPVLEY---POP-GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSK	127
DB	66	GDVEKQCORALIQHRPFIEWARPPQSGVEKAQAMPSPRLRTHLPTRLLPPSPFENNCK	125
QY	128	VIVYARNPKDLVSYOYPHRSLRTWSVRGTQFQFCRRFMNDKLGYGSGWFEHVQEFWEHRM	187
DB	126	FLYVARNVKDCMWSYVHFQRMNQVLPDPDGTWEEYFETETINGKAVMGSWFEHVKGWVEVKG	185
QY	188	DSNVLFLLKYEDMHDLDVTMWQLARFLGVCSDKAQLEALTEHC--HQLVDQ--CCNAEALP	244
DB	186	RYQILFLFYEDIKDKPKCEIRKVAQFMGKHLDVTLDKIVQETSPFKMKDPMINRSTVP	245
QY	245	-----VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF	280
DB	246	KSINQDISPPMRKGTGDMKNHFTVAQSHRLDELYRKMGEGVSI	292

RESULT 44  
 Q9XT99 RABIT PRELIMINARY; PRT; 292 AA.  
 AC Q9XT99;  
 DT 01-NOV-1999, integrated into UniProtKB/TrEMBL.  
 DT 01-NOV-1999, sequence version 1.  
 DT 07-FEB-2006, entry version 17.  
 DE Arylsulfotransferase StIA8.  
 OS *Oryctolagus cuniculus* (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;  
 OC *Oryctolagus*.  
 ON NCBI\_TaxId=9986;  
 RN [1]  
 RC NUCLEOTIDE SEQUENCE.  
 RX STRAIN=New Zealand White; TISSUE=Liver;  
 RX MEDLINE=2109889; PubMed=11181495;  
 RA Honma W., Kamiyama Y., Yoshinari K., Sasano H., Shimada M., Nagata K.,  
 RA Yamazoe Y.,  
 RT "Enzymatic characterization and interspecies difference of phenol  
 RT sulfotransferases, StIA forms.";  
 RL Drug Metab. Dispos. 29:274-281(2001).  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NonDerivs License  
 CC -----  
 CC EMBL; AB029494; BAA82295.1; -; mRNA.  
 DR HSSP; P50224; 1CUM.  
 DR SMR; Q9XT99; 8-289.  
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR InterPro; IPR000863; Sulfotransferase.  
 DR Pfam; PF00685; Sulfotransferase 1; 1.  
 DR ProDom; PD001218; Sulfotransferase; 1.  
 DR Transferase.  
 SEQ SEQUENCE 292 AA; 33841 MW; 2BF5F63A259DFFD1 CRC64;

Query Match	28.9%;	Score 442;	DB 2;	Length 292;
Best Local Similarity	35.3%;	Pred. No. 1.7e-28;		
Matches	95;	Conservative 50;	Mismatches 104;	Indels 20;
Gaps				

```
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC073295; AAH73295.1; -; mRNA.
DR GO; GO:0008146; F.sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
SQ SEQUENCE 297 AA; 34860 MW; 88D37337F60C6603 CRC64;

Query Match      28.8%; Score 440; DB 2; Length 297;
Best Local Similarity 32.9%; Pred. No. 2.6e-28;
Matches 97; Conservative 53; Mismatches 115; Indels 30; Gaps 7;

QY 13 GFSEKYVFEF-----HGVRLPFFCR---GKMBEIANFPVRSDVMIIVTPYKSGTSLLOQV 64
DB 13 GFSEKYVFEF-----HGVRLPFFCR---GKMBEIANFPVRSDVMIIVTPYKSGTSLLOQV 64
QY 65 VYLVSGADPDEIGLNMNIDQPLVLEY--POP---GLDIIKELTSRLIKSHLPYRFLPS 119
DB 65 VYLVSGADPDEIGLNMNIDQPLVLEY--POP---GLDIIKELTSRLIKSHLPYRFLPS 119
QY 59 VDSIMNAGDLKKVGRAPTHVRFPFLICNPSPVPCGVDVLEETQSPRIIKTHLPYELIPK 118
DB 59 VDSIMNAGDLKKVGRAPTHVRFPFLICNPSPVPCGVDVLEETQSPRIIKTHLPYELIPK 118
QY 120 DLHNGDSKVIYARNPKDLVSVYQFHSRLTMSYRGTFQFCRRFMNDKLGYSWFH 179
DB 120 DLHNGDSKVIYARNPKDLVSVYQFHSRLTMSYRGTFQFCRRFMNDKLGYSWFH 179
QY 119 SFWEHECKVIYARNADKNAVSYYYFDLMNKTQPHPTWEEYVGKFGNVPWGGFHHV 178
DB 119 SFWEHECKVIYARNADKNAVSYYYFDLMNKTQPHPTWEEYVGKFGNVPWGGFHHV 178
QY 180 QEFWEHRMDSNVLFLKYEDMRDLVTMVEQLARFLGVCDCQAQLEALTEHCH---QLVDQ 236
DB 180 QEFWEHRMDSNVLFLKYEDMRDLVTMVEQLARFLGVCDCQAQLEALTEHCH---QLVDQ 236
QY 179 IGWKKARAKHQIYVYEDMDKEDPKREIRKVRIFLGKELSDLEKICQHTSFKANKNP 238
DB 179 IGWKKARAKHQIYVYEDMDKEDPKREIRKVRIFLGKELSDLEKICQHTSFKANKNP 238
QY 237 CCNAEALP-----VGRGRVGLWKDIFTVSNMNEKFDLVYKQKMGKCDLTF 280
DB 237 CCNAEALP-----VGRGRVGLWKDIFTVSNMNEKFDLVYKQKMGKCDLTF 280
QY 239 MANYSAMPDVLDSQISPFMRKGEVSDWKVHFTVQQNEMFDAEYQKRMGTDLKF 293
DB 239 MANYSAMPDVLDSQISPFMRKGEVSDWKVHFTVQQNEMFDAEYQKRMGTDLKF 293

RESULT 46
Q3T0Y3 BOVIN PRELIMINARY; PRT; 296 AA.
AC Q3T0Y3;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DE 07-MAR-2006, entry version 4.
DE Hypothetical protein.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Crossbred x Angus; TISSUE=ileum;
RA Moore S.; Alexander L.; Brownstein M.; Guan L.; Lobo S.; Meng Y.;
RA Tanguchi M.; Wang Z.; Yu J.; Prange C.; Schreiber K.; Shenmen C.;
RA Wagner L.; Balu M.; Barabaz S.; Barber S.; Babakaiff R.; Beldand J.;
RA Chun E.; Del Rio L.; Gibson S.; Hanson R.; Kirkpatrick R.; Liu J.;
RA Matsuo C.; Mayo M.; Santos R.R.; Stott J.; Tsai M.; Wong D.;
RA Siddiqui A.; Holt R.; Jones S.J.; Marra M.A.;
RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC102208; AA102209.1; -; mRNA.
DR GO; GO:0008146; F.sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Hypothetical protein.
SQ SEQUENCE 296 AA; 34674 MW; 8A44653780078275 CRC64;

Query Match      28.5%; Score 436.5; DB 2; Length 296;
Best Local Similarity 33.9%; Pred. No. 5.1e-28;
```

```
Matches 94; Conservative 53; Mismatches 111; Indels 19; Gaps 4;

QY 23 HGVRLPFFCRGKMBEIANFPVRSDVMIIVTPYKSGTSLLOQVYLVSGADPDEIGLNMNI 82
DB 23 HGVRLPFFCRGKMBEIANFPVRSDVMIIVTPYKSGTSLLOQVYLVSGADPDEIGLNMNI 82
QY 16 HGCPITYAFANNWEKIEQFSRPPDDIMIVTPYKSGTTSWISVDVMDLHGDVCKCKROVI 75
DB 16 HGCPITYAFANNWEKIEQFSRPPDDIMIVTPYKSGTTSWISVDVMDLHGDVCKCKROVI 75
QY 83 DEQLPVLEYPOP-----GLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYARNPKD 137
DB 83 DEQLPVLEYPOP-----GLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYARNPKD 137
QY 76 TAKVPMLEALPGLRTSGLEQLEKNPSRVVKTHTLPDLIPKSFWENNCKIYIYARNAKD 135
DB 76 TAKVPMLEALPGLRTSGLEQLEKNPSRVVKTHTLPDLIPKSFWENNCKIYIYARNAKD 135
QY 138 LVSVYQFHSRLTMSYRGTFQFCRRFMNDKLGYSWFHVEHVEHVEHVEHVEHVEHVEH 197
DB 138 LVSVYQFHSRLTMSYRGTFQFCRRFMNDKLGYSWFHVEHVEHVEHVEHVEHVEHVEH 197
QY 136 VAVSFYHFDLMNQLPGLTGWGEVLEKLTGNVAGSFWHVKSWKKEGHPILFLFYE 195
DB 136 VAVSFYHFDLMNQLPGLTGWGEVLEKLTGNVAGSFWHVKSWKKEGHPILFLFYE 195
QY 198 DMHRDLVTMVEQLARFLGVCDCQAQLEALTEHC---HQLVDQ--CCNAEALP----- 244
DB 198 DMHRDLVTMVEQLARFLGVCDCQAQLEALTEHC---HQLVDQ--CCNAEALP----- 244
QY 196 DMKNPKQEKIKKVVRFLEKLNLDDELTKIYHTSFEMMKDNPVNYTHLPSEVMDHSS 255
DB 196 DMKNPKQEKIKKVVRFLEKLNLDDELTKIYHTSFEMMKDNPVNYTHLPSEVMDHSS 255
QY 245 -VGRGRVGLWKDIFTVSNMNEKFDLVYKQKMGKCDLTF 280
DB 245 -VGRGRVGLWKDIFTVSNMNEKFDLVYKQKMGKCDLTF 280
QY 256 FMRKGIAGDWKNYFTVAQNEKFDIYKEMSETELQF 292
DB 256 FMRKGIAGDWKNYFTVAQNEKFDIYKEMSETELQF 292

RESULT 47
STIE2_RAT STANDARD; PRT; 295 AA.
AC P52845;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-1996, sequence version 1.
DT 07-FEB-2006, entry version 38.
DE Estrogen sulfotransferase, isoform 2 (EC 2.8.2.4) (EST-2)
DE (Sulfotransferase, estrogen-prefering) (Estrone sulfotransferase).
GN Name=Ste2; Synonyms=Ste;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=Fischer 344; TISSUE=Liver;
RX MEDLINE=96305357; PubMed=8688469; DOI=10.1016/0167-4781(96)00065-6;
RA Rikke B.A.; Roy A.K.;
RT "Structural relationships among members of the mammalian
RT Sulfotransferase gene family.";
RL Biochim. Biophys. Acta 1307:331-338(1996).
CC -!- FUNCTION: Sulfation of estrone and estradiol. May control the
CC level of the estrogen receptor by sulfurylating free estradiol (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =
CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U50205; AA507681.1; -; mRNA.
DR HSP; P49891; IBO6.
DR SMR; P52845; 7-294.
DR Ensembl; ENSRNOG0000001957; Rattus norvegicus.
DR RGD; 3776; Ste.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Lipid-binding; Steroid-binding; Transferase.
FT CHAIN 1 295 Estrogen sulfotransferase, isoform 2.
FT FTID=PRO_0000085156.
FT NP_BIND 48 53 PAPS (By similarity).
FT NP_BIND 130 138 PAPS (By similarity).
FT NP_BIND 193 229 PAPS (By similarity).
FT NP_BIND 257 259 PAPS (By similarity).
```

```
FT ACT SITE 108 108 Proton acceptor (By similarity).
SQ SEQUENCE 295 AA; 35365 MW; 149B5C9D46039AAF CRC64;

Query Match 28.4%; Score 435; DB 1; Length 295;
Best Local Similarity 32.3%; Pred. No. 6.8e-28;
Matches 92; Conservative 60; Mismatches 115; Indels 18; Gaps 3;

QY 14 EFSEKYFEHGVRLPPFCRGKMEETANFPPRPSDVWIVTPYKSGTSLQEVVYLVSQAD 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 EYEVFGDFHGLMDKLTFTKYMEDVETFSARPDDLVTVPYKSGTGWIGEIVDMYKEGD 66
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 74 PDEIGLNMIDQLPVLVYQPQ---GLDIIKELTSPRLIKSHLPYRPLPSDLHNGDSKVI 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 67 VEKCKEDAIFNRIPLYECRNEDLINGIKQLKEKESPRIVKTHLPKALLPASFWKNCKII 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 130 YMARNPDLVSVYYQFHRSRLTMSYRGTFQBFRCRRFMNDKLGYSWFEHVOEFWEHRMDS 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 127 YLCRNAKDVVSVYYFFLLIKSYNPKNPSFSEFVEKFMEGQVPYGSWYDHVKSWEKSKNS 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 190 NVLFKVEDMRDLVTWVEQLARFLGVSCDKAQLAEALTEHC---HQLVDQCCNAEALP-- 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 187 RVLFMFYEDMKDIRREVVKLTIEFLERDPSAELVDRIIQTSTFQBMKNPCTNYSMPLPET 246
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 245 -----VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 247 MIDLKVSFPMRKGVGDWKNHPEALRERFEEHYQQQMKDCPVKF 291
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 48
ST186 RAT STANDARD; PRT; 295 AA.
AC P49890;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-1996, sequence version 1.
DT 07-FEB-2006, entry version 37.
DE Estrogen sulfotransferase, isoform 6 (EC 2.8.2.4) (EST-6)
DE (Sulfotransferase, estrogen-prefering) (Estrone sulfotransferase).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Liver;
RX MEDLINE=95161323; PubMed=7857871; DOI=10.1016/0960-0760(94)00147-E;
RA Falany J.L., Krasykh V., Mikheeva G., Falany C.N.;
RT "Isolation and expression of an isoform of rat estrogen
   sulfotransferase."
RL J. Steroid Biochem. Mol. Biol. 52:35-44(1995).
CC -!- FUNCTION: Sulfation of estrone and estradiol. May control the
   level of the estrogen receptor by sulfolating free estradiol.
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =
   adenosine 3',5'-bisphosphate + estrone 3-sulfate.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- DEVELOPMENTAL STAGE: Expressed in males rats, but not in females.
CC -!- INDUCTION: Induced by androgens and suppressed by estrogens.
CC Expression is under the influence of pituitary growth hormone and
   thyroid hormone.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC -----
CC Coprighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC
DR EMBL; S76490; AAB33442.1; -; mRNA.
DR HSP; P49891; 1B06.
DR SMR; P49890; 7-294.
DR Ensembl; ENSRNOG00000001957; Rattus norvegicus.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Lipid-binding; Steroid-binding; Transferase.
```

```
FT CHAIN 1 295 Estrogen sulfotransferase, isoform 6.
FT FTID=PRO_0000085158.
FT NP_BIND 48 53 PAPS (By similarity).
FT NP_BIND 130 138 PAPS (By similarity).
FT NP_BIND 229 229 PAPS (By similarity).
FT NP_BIND 257 259 PAPS (By similarity).
FT ACT_SITE 108 108 Proton acceptor (By similarity).
SQ SEQUENCE 295 AA; 35303 MW; 9FBD5861AFDC9B05 CRC64;

Query Match 28.4%; Score 435; DB 1; Length 295;
Best Local Similarity 31.8%; Pred. No. 6.8e-28;
Matches 93; Conservative 62; Mismatches 105; Indels 32; Gaps 4;

QY 14 EFSEKYFEHGVRLPPFCRGKMEIANFPPRPSDVWIVTPYKSGTSLQEVVYLVSQAD 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 EYEVFGDFHGLVDKLTFTKYMEDVETFSARPDDLVTVPYKSGTGWIGEIVDMYKEGD 66
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 74 PDEIGLNMIDQLPVLVYQPQ---GLDIIKELTSPRLIKSHLPYRPLPSDLHNGDSKVI 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 67 VEKCKEDAIFNRIPLYECRNEDLINGIKQLKEKESPRIVKTHLPKALLPASFWKNCKII 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 130 YMARNPDLVSVYYQFHRSRLTMSYRGTFQBFRCRRFMNDKLGYSWFEHVOEFWEHRMDS 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 127 YLCRNAKDVVSVYYFFLLIKSYNPKNPSFSEFVEKFMEGQVPYGSWYDHVKSWEKSKNS 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 190 NVLFKVEDMRDLVTWVEQLARFLGVSCDKAQLAEALTEHCQLV-----DQCCN 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 187 RVLFMFYEDMKDIRREVVKLTIEFL-----ERDPLAELVDKIIQTSTFQBMKNPCTN 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 240 AEALP-----VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 YSMLPETMIDLKVSFPMRKGVGDWNRHFPALRERFEEHYQRHMKDCPVTF 291
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 49
ST181 BRARE STANDARD; PRT; 299 AA.
AC Q6PH37; Q7T1C8;
DT 31-AUG-2004, integrated into UniProtKB/Swiss-Prot.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 15.
DE Cytosolic sulfotransferase 1 (EC 2.8.2.-) (SULT1 ST1).
GN Name=sultiat1;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], ACTIVITY, TISSUE SPECIFICITY, AND ENZYME
   REGULATION.
RX MEDLINE=22640919; PubMed=12755695;
RX DOI=10.1046/j.1432-1033.2003.03608.x;
RA Sugahara T., Liu C.-C., Pai T.G., Colloidi P., Suiko M., Sakakibara Y.,
RA Nishiyama K., Liu M.-C.;
RT "Sulfation of hydroxychlorobiphenyls. Molecular cloning, expression,
   and functional characterization of zebrafish SULT1
   sulfotransferases."
RL Eur. J. Biochem. 270:2404-2411(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Kidney;
RG NIH - Zebrafish Gene Collection (ZGC) project;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Catalyzes the conjugation of sulfate to a variety of
   xenobiotic and endogenous compounds, including 2-naphthol,
   hydroxychlorobiphenyls, dopamine and T3 (triiodo-L-thyronine).
CC -!- ENZYME REGULATION: Inhibited by Co(2+), Zn(2+), Cd(2+) and Pb(2+)
   ions. Inactivated by Hg(2+) and Cu(2+) ions.
CC -!- BIOPHYSICOCHEMICAL PROPERTIES:
   pH dependence:
   Optimum pH is 6.0-9.0;
   Temperature dependence:
```

```

CC      Thermostable from 20 to 43 degrees Celsius;
CC      SUBCELLULAR LOCATION: Cytoplasm.
CC      TISSUE SPECIFICITY: Expressed in liver.
CC      SIMILARITY: Belongs to the sulfotransferase family.
CC      -----
CC      Copylefted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      -----
DR      EMBL; AY181064; AA064983.1; -; mRNA.
DR      EMBL; BC056729; AAH56729.1; -; mRNA.
DR      ZFIN; ZDB-GENE-030131.2144; sult1t1.
DR      GO; GO:0005737; C:cytoplasm; NAS.
DR      GO; GO:0008146; P:sulfotransferase activity; IDA.
DR      GO; GO:0006805; P:xenobiotic metabolism; IDA.
DR      InterPro; IPR000863; Sulfotransferase.
DR      Pfam; PF00685; Sulfotransferase.
DR      ProDom; PD001218; Sulfotransferase; 1.
KW      Catecholamine metabolism; Transferase.
FT      CHAIN   1 299   Cytosolic sulfotransferase 1.
FT              /FTid=PRO_0000085173.
FT      NP_BIND 51    56    PAPS (By similarity).
FT      NP_BIND 135   143   PAPS (By similarity).
FT      NP_BIND 199   235   PAPS (By similarity).
FT      NP_BIND 261   263   PAPS (By similarity).
FT      ACT_SITE 113  113   Proton acceptor (By similarity).
FT      CONFLICT 3    3    I -> M (in Ref. 1).
FT      CONFLICT 193  193   N -> T (in Ref. 1).
SQ      SEQUENCE 299 AA; 35035 MW; 805D82915B45EB6A CRC64;
      Query Match      28.4%; Score 434.5; DB 1; Length 299;
      Best Local Similarity 36.1%; Pred. No. 7.6e-28;
      Matches 101; Conservative 51; Mismatches 109; Indels 19; Gaps 6;
OY      20  FEFGVRLPPFCRGKMEIANFPVPDVIYTPKSGTSLQEVYLVLSQAGD-PDEIG 78
Db      16  FEFGISMINHPFENWEKVNQFQRPDDILIAITYPKAGITWYSYILDLLYFCENAPEHT 75
OY      79  LWNIDEQLPVLE-----YQPPGLDIKEL-TSPRLKSHLPYRFLPSDLHNGDSKVIYVAR 133
Db      76  SQPIYMYRPFLESCFKVIASGTGLADNMITSRLIKTHLPVLQIPKSFWEQNSRVYVVAR 135
OY      134  NPKDLVSVYOPHRSIRLTSYRGTFQECRRMNDKLGVSFHFVQSFWEHRMD-SNVL 192
Db      136  NAKQNVSVYFHRDNIIVPEPDGWNTELRFMQSGKVFPGPDYHVGWKEKKQYTNLL 195
OY      193  FLKYEDMRDLVTWVEQLARFLGVSCKQALEALTEHCH--QLVDQCCNAELPV---- 245
Db      196  YLFYEDLVEDTGREVDRLCSFLGLSTSVSDREKITKDQVFDAMKQNKNTYSTLPVMDPK 255
OY      246  -----GGRGVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
Db      256  ISPFMRKGVGDGKNHFTVAQNEQPDVEYKQKKNATVKF 295
RESULT 50
STIE1_HUMAN STANDARD; PRT; 294 AA.
AC P49888; Q9N6X5;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-1996, sequence version 1.
DT 07-FEB-2006, entry version 48.
DE Estrogen sulfotransferase (EC 2.8.2.4) (Sulfotransferase, estrogen-
DE preferring) (EST-1).
GN Name=SULT1E1; Synonyms=STE;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=94242031; PubMed=8185618;

```

```

RA Aksoy I.A., Wood T.C., Weinshilboum R.;
RT "Human liver estrogen sulfotransferase: identification by cDNA cloning
RT and expression.";
RL Biochem. Biophys. Res. Commun. 200:1621-1629(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=96079087; PubMed=8530066;
RA Her C., Aksoy I.A., Kimura S., Brandriff B.F., Wasmuth J.J.,
RA Weinshilboum R.M.;
RT "Human estrogen sulfotransferase gene (STE): cloning, structure, and
RT chromosomal localization.";
RL Genomics 29:16-23(1995).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=95298714; PubMed=7779757; DOI=10.1016/0960-0760(95)00015-R;
RA Falany C.N., Krasnykh V., Falany J.L.;
RT "Bacterial expression and characterization of a cDNA for human liver
RT estrogen sulfotransferase.";
RL J. Steroid Biochem. Mol. Biol. 52:529-539(1995).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20009807; PubMed=10541560; DOI=10.1093/molehr/5.11.995;
RA Rubin G.L., Harrold A.J., Mills J.A., Falany C.N., Coughtrie M.W.H.;
RT "Regulation of sulphotransferase expression in the endometrium during
RT the menstrual cycle, by oral contraceptives and during early
RT pregnancy.";
RL Mol. Hum. Reprod. 5:995-1002(1999).
RN [5]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANT TYR-22.
RA Rieder M.J., Livingston R.J., Daniels M.R., Chung M.-W.,
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA Shackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RT "NIHES-SNPs, environmental genome project, NIHES ES15478, Department
RT of Genome Sciences, Seattle, WA (URL: http://egp.gs.washington.edu).";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [7]
RP NUCLEOTIDE SEQUENCE OF 182-294.
RC TISSUE=Liver, and Spleen;
RA Her C., Szumlanaki C., Aksoy I., Weinshilboum R.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF MUTANT VAL-269 IN COMPLEX
RP WITH 3'-PHOSPHATE-ADENOSINE-5'-PHOSPHATE SULFATE (PAPS), AND
RP MUTAGENESIS OF SER-137.
RX PubMed=11884392; DOI=10.1074/jbc.M111651200;
RA Pedersen L.C., Petrotchenko E., Shevtsov S., Negishi M.;
RT "Crystal structure of the human estrogen sulfotransferase-PAPS

```

RT complex: evidence for catalytic role of Ser137 in the sulfonyl  
 RT transfer reaction.";  
 RL J. Biol. Chem. 277:17928-17932(2002).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) IN COMPLEX WITH ADENOSINE  
 RP 3',5'-BISPHOSPHATE (PAP) AND A HYDROXYLATED POLYCHLORINATED BIPHENYL  
 RP (OH-PCB).  
 RX PubMed:12782487;  
 RA Shevtsov S., Petrotchenko E.V., Pedersen L.C., Negishi M.;  
 RT "Crystallographic analysis of a hydroxylated polychlorinated biphenyl  
 RT (OH-PCB) bound to the catalytic estrogen binding site of human  
 RT estrogen sulfotransferase.";  
 RL Environ. Health Perspect. 111:884-888(2003).  
 CC -|- FUNCTION: May control the level of the estrogen receptor by  
 CC sulfonylating free estradiol. Maximally sulfates beta-estradiol  
 CC and estrone at concentrations of 20 nM. Also sulfates  
 CC dihydroepiandrosterone, pregnenolone, ethinylestradiol, equalenin,  
 CC diethylstilbestrol and 1-naphthol, at significantly higher  
 CC concentrations; however, cortisol, testosterone and dopamine are  
 CC not sulfated.  
 CC -|- CATALYTIC ACTIVITY: 3'-phosphoadenyl sulfate + estrone =  
 CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.  
 CC -|- SUBUNIT: Homodimer.  
 CC -|- SUBCELLULAR LOCATION: Cytoplasm.  
 CC -|- TISSUE SPECIFICITY: Liver, intestine and at lower level in the  
 CC kidney.  
 CC -|- SIMILARITY: Belongs to the sulfotransferase family.  
 CC  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC  
 CC EMBL: U08098; AA82125.1; -; mRNA.  
 CC EMBL: U20521; AAC50286.1; -; Genomic\_DNA.  
 CC EMBL: U20515; AAC50286.1; JOINED; Genomic\_DNA.  
 CC EMBL: U20516; AAC50286.1; JOINED; Genomic\_DNA.  
 CC EMBL: U20517; AAC50286.1; JOINED; Genomic\_DNA.  
 CC EMBL: U20518; AAC50286.1; JOINED; Genomic\_DNA.  
 CC EMBL: U20519; AAC50286.1; JOINED; Genomic\_DNA.  
 CC EMBL: U20520; AAC50286.1; JOINED; Genomic\_DNA.  
 CC EMBL: S77383; AA334601.1; -; mRNA.  
 CC EMBL: Y11195; CAA72079.1; -; mRNA.  
 CC EMBL: AY436634; AAQ97179.1; -; Genomic\_DNA.  
 CC EMBL: BC027956; AAH27956.1; -; mRNA.  
 CC EMBL: U55764; AA951658.1; -; mRNA.  
 CC F01; J02229; J02229.  
 CC PDB: 1G3M; X-ray; A/B=1-294.  
 CC PDB: 1HY3; X-ray; A/B=1-294.  
 CC IntAct; P49888; -;  
 CC EMBL: ENSG00000109193; Homo sapiens.  
 CC H-InvDB; HIX0022597; -;  
 CC HGNC; HGNC:11377; SULT1E1.  
 CC MIM; 600043; Gene.  
 CC GO; GO:0004304; F:estrogen sulfotransferase activity; TAS.  
 CC GO; GO:0008202; P:steroid metabolism; TAS.  
 CC InterPro; IPR000863; Sulfotransferase.  
 CC Pfam; PF00685; Sulfotransferase\_1; 1.  
 CC ProDom; PD001218; Sulfotransferase; 1.  
 CC 3D-structure; Lipid-binding; Polymorphism; Steroid-binding;  
 KW Transferase.  
 CHAIN 1 294 Estrogen sulfotransferase.  
 FT FTID=PRO\_0000085153.  
 FT NP\_BIND 47 52 PAPS.  
 FT NP\_BIND 129 137 PAPS.  
 FT NP\_BIND 192 228 PAPS.  
 FT NP\_BIND 256 258 PAPS.  
 FT ACT\_SITE 107 107 Proton acceptor (By similarity).  
 FT VARIANT 22 22 D -> Y (in dbSNP:11569705).  
 FT FTID=VAR\_018907.  
 FT MUTAGEN 137 137 S->A: Decreased gradually the catalytic  
 FT activity.  
 FT MUTAGEN 137 137 S->C: Decreased gradually the catalytic  
 FT activity.  
 FT MUTAGEN 269 269 V->E: Does not prevent the formation of

FT CONFLICT 154 154 homodimer.  
 FT HELIX 4 10 F -> L (in Ref. 6).  
 FT STRAND 11 14  
 FT TURN 15 16  
 FT STRAND 17 20  
 FT HELIX 21 25  
 FT TURN 26 26  
 FT HELIX 27 31  
 FT TURN 32 32  
 FT TURN 37 38  
 FT STRAND 40 45  
 FT TURN 46 47  
 FT STRAND 48 49  
 FT HELIX 50 61  
 FT TURN 62 64  
 FT HELIX 66 69  
 FT STRAND 70 72  
 FT HELIX 74 77  
 FT STRAND 78 78  
 FT STRAND 80 80  
 FT TURN 81 82  
 FT STRAND 84 85  
 FT TURN 86 88  
 FT STRAND 89 90  
 FT HELIX 91 96  
 FT TURN 97 97  
 FT STRAND 100 101  
 FT STRAND 103 106  
 FT HELIX 110 112  
 FT STRAND 113 113  
 FT HELIX 115 119  
 FT TURN 120 121  
 FT STRAND 123 128  
 FT HELIX 131 144  
 FT STRAND 145 145  
 FT TURN 146 147  
 Query Match 28.4%; Score 434; DB 1; Length 294;  
 Best Local Similarity 33.3%; Pred. No. 8.2e-28;  
 Matches 95; Conservative 56; Mismatches 116; Indels 18; Gaps 3;  
 QY 14 EFESKYPFHGVRLLPPFCRCGMIEIANFPVPSPDVWVTPKSGTSLIQEYVYLVSQAD 73  
 Db 6 DYKEFEVHGILMYKDFVKYWDNVEAFQARPDDLVIATYKSGTTWVSEIVYMYKBDG 65  
 QY 74 PDEIGLMNIDQLPVLEYQPQ-----GLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVI 129  
 Db 66 VEKCKEDVIFNRIPFLCCKENLMNGVKQLDENMSPRIVKTHLPPELLPASFEKDKCKII 125  
 QY 130 YMABNPKDLVYVYQFHRSLTMSYRGTFQFCRRFMDNKLGVSWFPEHVQSFWEHRMDS 189  
 Db 126 YLCRNKADVAVSFYFFFLMVAHGNPQSPFPEFVKFGQGVPGYSGWKVKSMMWKGKSP 185  
 QY 190 NVLFKLYEDMDHRLVTWVEQLARFLGVSCDKAQLAEALTEHCH-----QLVDQ 236  
 Db 186 RVLFELFYEDLKEDIRKVKILHFLERKPSSELDVRIIHHTSFQEMKNPNSTNYTLLPDE 245  
 QY 237 CCNAEALP-VGRGRVGLMKDIFTVSMNEKFDLVYKQMGKCDLTF 280  
 Db 246 IMNQKLSPPMRKGITGDWKNHFTVALNEKFDKHYEQQMKESTLKF 290  
 RESULT 51  
 Q53X91 HUMAN PRELIMINARY; PRT; 294 AA.  
 ID Q53X91 HUMAN  
 AC Q53X91;  
 DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.  
 DT 24-MAY-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 3.  
 DE SULT1E1 protein (fragment).  
 GN Name=SULT1E1;  
 OS Homo sapiens (Human).



[illegible]



```
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Teijima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AK015750; BAB29956.1; -; mRNA.
DR HSSP; P49891; 1B06.
DR SMR; Q9D566; 7-294.
DR Ensemble; ENSMUSG0000029272; Mus musculus.
DR MGI; MGI:98431; Sult1el.
DR GO; GO:0008210; P:estrogen metabolism; IMP.
DR GO; GO:0007565; P:pregnancy; IMP.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 295 AA; 35576 MW; 5526B3FC362BFB0E CRC64;

Query Match 28.2%; Score 432; DB 2; Length 295;
Best Local Similarity 33.0%; Pred. No. 1.2e-27;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYFEHGVRLPPFCRGKMEETANFPVRSDWIVTPYKSGTSLQEVVYLVSQAGD 73
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
7 EYEVFGFGRGLVMDKRFYKWEDEMFLARDDIVATPKSGTWTSEVYMYIKEGD 66

QY 74 PDEIGLWNIDQLPVLEYPQP---GLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKV 129
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
67 VEKKEDAIENRIPVLECRNEDLINGIQLKKEKESFRIVKTHLPVQLVPLFWYKCKV 126

QY 130 YMARNPDLVSYQYFHRSLRTMSYRGTFQFCRPFMNDKLGYSWFHFVQEFWEHRMDS 189
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
127 YLCRNAKDVAVSYYFFLLMITSYPNPKSFSEFVEKFMQGVPGYSWYDHVRAWEKSKNS 186

QY 190 NVFLFKYEDMHRDLVTWVEQLARFLGVSCDRAQLEALTEHCH-----QLVDQ 236
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
187 RVLFMFYEDMKEDIRREVVKLIEFLERKPSAELVDRIVQHTSFQEMKNPNSTNYTMDEE 246

QY 237 CCNAEALP-VGRGRVGLMKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
247 MNQKVSFPMRKGIIIGDWKNHFPPEALRERFDEHYKQKMKDCTVKF 291

RESULT 54
Q5HZV7_XENTR PRELIMINARY; PRT; 287 AA.
ID Q5HZV7_XENTR
AC Q5HZV7
DT 15-FEB-2005, integrated into UniProtKB/TrEMBL.
DT 15-FEB-2005, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE Sulfotransferase family, cytosolic, 1A, phenol-preferring, member 1.
GN Name=sult1al1; prov;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=83364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole body;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
```

```
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole body;
RA Klein S., Gerhard D.S.;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC088870; AAH88870.1; -; mRNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 287 AA; 33894 MW; 89CFB2DC7991479B CRC64;

Query Match 28.2%; Score 431.5; DB 2; Length 287;
Best Local Similarity 32.6%; Pred. No. 1.3e-27;
Matches 94; Conservative 58; Mismatches 115; Indels 21; Gaps 4;

QY 12 PGEPESKYFEHGVRLPPFCRGKMEETANFPVRSDWIVTPYKSGTSLQEVVYLVSQG 71
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2 PGKIQI-----VEGVSIAEDIASNQIQSQFQARPGDVLVIATPKAGTTWQEIVDLIINE 57

QY 72 ADPDEIGLWNIDQLP---VLEYPQGLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKV 128
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
58 GNEEICRRSPHTHERMPFVEVLHMMKPGPEEVNAMPSPVLKTHLPVQLVPLFWYKCKV 117

QY 129 IYMARNPDLVSYQYFHRSLRTMSYRGTFQFCRPFMNDKLGYSWFHFVQEFWEHRMD 188
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
118 IYVARNPRDVTVSYYFDHTITFHPAPGSWEELHFRFMKGDVGWGSYDHVKFGEWQKQ 177

QY 189 SNVLFLKYEDMHRDLVTWVEQLARFLGVSCDRAQLEA---LTEHCHQLVDQCCNAEALP- 244
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
178 HNLIYLFYEDIKQNPPIHRIKRWRFDKDLSEEVLEKIVHLSSPFDHMKDNPMANFSAPPS 237

QY 245 -----VGRGRVGLMKDIFTVSMNEKFDLVYKQKMGKCDLTFDF 282
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
238 DVVDQSQYKFMKRGKVGDKWSKSHFTVQQNEMFEKYQQQMHGSAMKFRY 285

RESULT 55
ST1E1_MOUSE
ID ST1E1_MOUSE STANDARD; PRT; 295 AA.
AC P49891;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
```

DT 15-JUL-1999, sequence version 2.  
 DT 07-PEB-2006, entry version 47.  
 DE Estrogen sulfotransferase, testis isoform (EC 2.8.2.4)  
 DE (Sulfotransferase, estrogen-prefering).  
 GN Name=Sult1e1; Synonyms=Ste;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [mRNA].  
 RC STRAIN=C57BL/KSJ-DB/DB; TISSUE=Testis;  
 RX MEDLINE=95269690; PubMed=7750469; DOI=10.1210/en.136.6.2477;  
 RA Song W.-C., Moore R., McEachlan J.A., Negishi M.;  
 RT "Molecular characterization of a testis-specific estrogen  
 sulfotransferase and aberrant liver expression in obese and  
 diabetic C57BL/KSJ-db/db mice.";  
 RL Endocrinology 136:2477-2484(1995).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) IN COMPLEX WITH ADENOSINE  
 3',5'-BISPHOSPHATE (PAP), AND SEQUENCE REVISION TO 113.  
 RC STRAIN=C57BL/KSJ-DB/DB; TISSUE=Testis;  
 RX MEDLINE=98025056; PubMed=9360604;  
 RA Kakuta Y., Pedersen L.G., Carter C.W., Negishi M., Pedersen L.C.;  
 RT "Crystal structure of estrogen sulphotransferase.";  
 RL Nat. Struct. Biol. 4:904-908(1997).  
 CC -!- FUNCTION: Sulfation of estradiol and estrone. May control the  
 level of the estrogen receptor by sulphylation free estradiol.  
 CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =  
 adenosine 3',5'-bisphosphate + estrone 3-sulfate.  
 CC -!- SUBUNIT: Homodimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasm.  
 CC -!- TISSUE SPECIFICITY: Testis and at very low level in the placenta.  
 CC -!- MISCELLANEOUS: Abnormal high expression in liver in obese and  
 diabetic C57BL/KSJ-DB/DB strain mice. Female > male. Normal  
 level in liver.  
 CC -!- SIMILARITY: Belongs to the sulfotransferase family.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; S78182; AAB34320.1; -; mRNA.  
 DR PDB; 1AQU; X-ray; A/B=1-295.  
 DR PDB; 1AQY; X-ray; A/B=1-295.  
 DR PDB; 1BO6; X-ray; A/B=1-295.  
 DR PDB; 1BO6; X-ray; A/B=1-295.  
 DR IntAct; P49891; -;  
 DR Ensembl; ENSMUSG0000029272; Mus musculus.  
 DR MGI; MGI:98431; Sult1e1.  
 DR LinkHub; P49891; -;  
 DR GO; GO:0008210; P:estrogen metabolism; IMP.  
 DR GO; GO:0007565; P:pregnancy; IMP.  
 DR InterPro; IPR000863; Sulfotransferase.  
 DR Pfam; PF00685; Sulfotransferase 1; 1.  
 DR ProDom; PD001218; Sulfotransferase; 1.  
 KW 3D-structure; Lipid-binding; Steroid-binding; Transferase.  
 FT CHAIN 1 295  
 FT Estrogen sulfotransferase, testis  
 FT Isoform.  
 FT NP\_BIND 48 53  
 FT NP\_BIND 130 138  
 FT NP\_BIND 193 229  
 FT NP\_BIND 257 259  
 FT ACT\_SITE 108 108  
 FT CONFLICT 113 113  
 FT HELIX 8 11  
 FT STRAND 12 15  
 FT TURN 16 17  
 FT STRAND 18 21  
 FT HELIX 22 24  
 FT TURN 25 27  
 FT HELIX 28 32  
 FT TURN 33 33

FT	TURN	38	39
FT	STRAND	41	45
FT	TURN	47	48
FT	STRAND	49	50
FT	HELIX	51	62
FT	TURN	63	64
FT	HELIX	75	78
FT	STRAND	79	79
FT	STRAND	81	81
FT	TURN	82	83
FT	STRAND	85	86
FT	TURN	87	89
FT	STRAND	90	91
FT	HELIX	92	97
FT	TURN	98	98
FT	STRAND	101	102
FT	STRAND	104	107
FT	HELIX	111	113
FT	STRAND	114	114
FT	HELIX	116	120
FT	TURN	121	122
FT	STRAND	124	129
FT	HELIX	132	145
FT	STRAND	146	146
FT	TURN	147	148
FT	STRAND	149	149
FT	STRAND	153	153
FT	HELIX	155	164
FT	TURN	165	165
FT	STRAND	167	167
FT	TURN	168	169
FT	HELIX	172	182
FT	TURN	183	184
FT	STRAND	185	185
FT	TURN	186	187
FT	STRAND	188	192
FT	HELIX	193	198
FT	TURN	200	210
FT	HELIX	211	212
FT	STRAND	217	226
FT	TURN	227	227
FT	STRAND	228	228
FT	HELIX	229	234
FT	TURN	236	242
FT	STRAND	243	243
FT	TURN	245	247
FT	STRAND	248	248
FT	TURN	250	252
FT	STRAND	253	253
FT	STRAND	255	256
FT	STRAND	261	261
FT	HELIX	263	266
FT	TURN	267	267
FT	HELIX	270	284
FT	TURN	285	286
SQ	SEQUENCE	295 AA;	35590 MW;

Query Match 28.2%; Score 431; DB 1; Length 295;  
 Best Local Similarity 33.0%; Pred. No. 1.5e-27;  
 Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

QY	14	EFESKYEFHGVRLPPFCRGKWEIAPPPVPSDVIWTVYKSGTSLAQEVVYLVSQAD	73
DB	7	EYVEFGFGRGLMDKRFKTKYWDVEMFLAPDDLVVIATYPASGTTWSEVVYMIYKGD	66
QY	74	PDEIGLMNIDQLPVLEYPQP-GLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVI	129
DB	67	VEKCKEDAIENRIPVLECRNEDLINGIKLQKESPRIVKTHLPKLLPASFEWKNCKMI	126
QY	130	YMAENPKDLVSYVYQFHRSRLTMSYRGTFQFCRRFMDKLGYSWFHVEHVEHMDS	189
DB	127	YLCRNAKDVAVSYYFFLLIMITSYPNPKSFSEFVKFMQGPYPYGSWDHVKAWNEKSKNS	186

## RESULT 56

RA SClausberg K.;  
PI submitted (JUN-2002) to the EMPL/ComBank/DBPT databases

DR HSSP; P49891; 1AQY.

RA Stapleton M., Soares M.B., Bernaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Kana S.S., Loquerriano N.A., Peters C.O., Atkinson K.B., Murray S.B.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

FAHEY J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences." ;

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).





```
QY 6 AETPSTGCEFFSKYFPHGVRLPPFCRCGMEEIANFPVRPSDVTWITVPKSGTSLLOQVW 65
Db 2 AONFSNMEPLKPLVPVKGIPLIKFAETMQLQNFMTAPDDVLITVPKSGTTWMSIM 61
QY 66 YLVSGQADPDEIGLMNIDQLPVLVEYP-----QPGLDIIKELTSPRLIKSHLPFLPSDL 121
Db 62 DMIVGGKLDKCGRPVYARIPFLFSCPGVPGLLETKETAPRIKTHLPKSLPQSL 121
QY 122 HNGDSKVITYMARNPKDLVSVYQFHRSLRTWSYRGTFQRCRRFNMNDKLGVSFWFHVQE 181
Db 122 LDQIKVITYVARNAKDVVSYNFYKMAKHPDPTGWTESFLENFMGDKVSGVSYQHVKE 181
QY 182 FWEHRMDSNVFLKVEDMHRDLVTWVEQLARFLGVSCDKAOLEALTEH 229
Db 182 MWELRTHPVLYLFYEDMKENPKREIKKILEFLGRSLPEETVDLIVHH 229
RESULT 60
ST1E3 RAT STANDARD; PRT; 295 AA.
AC P49889; Q9QWS0;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-1996, sequence version 1.
DT 07-FEB-2006, entry version 41.
DE Estrogen sulfotransferase, isoform 3 (EC 2.8.2.4) (EST-3)
DE (Sulfotransferase, estrogen-prefering) (Estrone sulfotransferase).
GN Name=Ste;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 258-265.
RC TISSUE=Liver;
RX MEDLINE=92261615; PubMed=1374839; DOI=10.1210/me.6.4.589;
RA Demyan W.P., Song C.S., Kim D.S., Her S., Gallwitz W., Rao T.R.,
RA Slomczynska M., Chatterjee B., Roy A.K.;
RT "Estrogen sulfotransferase of the rat liver: complementary DNA cloning
RT and age- and sex-specific regulation of messenger RNA.";
RL Mol. Endocrinol. 6:589-597(1992).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=95161323; PubMed=7857871; DOI=10.1016/0960-0760(94)00147-E;
RA Falany J.L., Kraenkh V., Mikhieva G., Falany C.N.;
RT "Isolation and expression of an isoform of rat estrogen
RT sulfotransferase.";
RL J. Steroid Biochem. Mol. Biol. 52:35-44(1995).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Astapova I.I., Smirnov A.N., Rubtsov P.M.;
RT "PCR amplification and structural analysis of two paralogous rat
RT estrogen sulfotransferase genes.";
RL Mol. Biol. (Mosk.) 36:635-642(2002).
CC -!- FUNCTION: Sulfation of estrone and estradiol. May control the
CC level of the estrogen receptor by sulfonylating free estradiol.
CC -!- CATALYTIC ACTIVITY: 3'-phosphadenylyl sulfate + estrone =
CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- TISSUE SPECIFICITY: Liver of young mature males and uterus.
CC -!- DEVELOPMENTAL STAGE: Expressed only in the liver of young adult
CC animals (100 days old) and is absent in the prepubertal male (27
CC days old), senescent male (800 days old) and female liver.
CC -!- INDUCTION: Induced by androgens and suppressed by estrogens. The
CC expression is under the influence of pituitary growth hormone and
CC thyroid hormone. Is regulated by progesterone in the uterus.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Distributed under the UniProt Consortium, see http://www.uniprot.org/terms
CC Copyrighted by the UniProt Consortium, Attribution-NoDerivs License
```

```
CC -----
DR EMBL; M86758; AAA41128.1; -; mRNA.
DR EMBL; S76489; AAB33441.1; -; mRNA.
DR EMBL; AJ131835; CAAL0515.2; -; Genomic DNA.
DR EMBL; AJ298109; CAAL0515.2; JOINED; Genomic DNA.
DR EMBL; AJ298110; CAAL0515.2; JOINED; Genomic DNA.
DR EMBL; AJ298111; CAAL0515.2; JOINED; Genomic DNA.
DR EMBL; AJ298112; CAAL0515.2; JOINED; Genomic DNA.
DR EMBL; AJ298113; CAAL0515.2; JOINED; Genomic DNA.
DR EMBL; AJ298114; CAAL0515.2; JOINED; Genomic DNA.
DR PIR; A41930; A41930.
DR HSSP; P49891; 1AQU.
DR SMR; P49889; 7-294.
DR Ensembl; ENSRNOG0000001957; Rattus norvegicus.
DR RGD; 3776; Ste.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Direct protein sequencing; Lipid-binding; Steroid-binding;
KW Transferase.
FT CHAIN 1 295 Estrogen sulfotransferase, isoform 3.
FT NP_BIND 48 53 /FTID=PRO_0000085157.
FT NP_BIND 130 138 PAPS (By similarity).
FT NP_BIND 193 229 PAPS (By similarity).
FT NP_BIND 257 259 PAPS (By similarity).
FT ACT_SITE 108 108 Proton acceptor (By similarity).
FT CONFLICT 150 150 P -> Q (in Ref. 2).
FT CONFLICT 238 238 T -> I (in Ref. 2).
FT CONFLICT 295 295 L -> P (in Ref. 2).
SQ SEQUENCE 295 AA; 35416 MW; A77807A21DD2E7EB CRC64;
Query Match 28.0%; Score 429; DB 1; Length 295;
Best Local Similarity 32.3%; Pred. No. 2.2e-2;
Matches 92; Conservative 60; Mismatches 115; Indels 18; Gaps 3;
QY 14 EFESKYFEHGVRLPPFCRCGMEEIANFPVRPSDVTWITVPKSGTSLLOQVYLVVLSQAD 73
Db 7 EYDVFQDFHGLMDKRFYKWEVTEFLARPDDLIVTPKSGSTWSEIVDMYIKGD 66
QY 74 PDEIGLMNIDQLPVLVEYPQ-----GLDIIKELTSPRLIKSHLPFLPSDLHNGDSKVI 129
Db 67 VEKCKEDALFNRIPOLECRNEDLINGIKLQKESPRIVKTHLPKLLPASFWKNCXII 126
QY 130 YMAENPKDLVSVYQFHRSLRTWSYRGTFQRCRRFNMNDKLGVSFWFHVQEFWHRMDS 189
Db 127 YLCRNKADVVVSYFYFFFLIMKSYNPNSFSFVEKFMEGQVPYGSWYDHVKSWMKSKNS 186
QY 190 NVLFLKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHC---HQLVDQCCNAEALP-- 244
Db 187 RVLFMFYEDMKEDIRREVVKLIEFLERDPSAELVDRIIQTISFQEMKNPCTNYSMLPET 246
QY 245 -----VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 247 MDLKVSPFMKRGIVGDKWKNHFPALRERFEEHYQQQMKDCPVKF 291
RESULT 61
Q3T0S9 BOVIN
ID Q3T0S9_BOVIN PRELIMINARY; PRT; 295 AA.
AC Q3T0S9;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DT 07-MAR-2006, entry version 4.
DE Hypothetical protein.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Crossbred x Angus; TISSUE=ileum;
```



RA Moore S., Alexander L., Brownstein M., Guan L., Lobo S., Meng Y.,  
 RA Tanguchi M., Wang Z., Yu J., Frange C., Schreiber K., Shenmen C.,  
 RA Wagner L., Bala M., Barbazul S., Barber S., Babakaiff R., Beland J.,  
 RA Chun E., Del Rio L., Gibson S., Hanson R., Kirkpatrick R., Liu J.,  
 RA Matsuo C., Mayo M., Santos R.R., Scott J., Tsai M., Wong D.,  
 RA Siddiqui A., Holt R., Jones S.J., Marra M.A.;  
 RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.  
 CC  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC  
 CC EMBL: BC102274; AI02275.1; -; mRNA.  
 DR GO: GO:0008146; F:sulfotransferase activity; IEA.  
 DR InterPro: IPR000863; Sulfotransferase.  
 DR Pfam: PF00685; Sulfotransferase\_1.  
 DR ProDom: PD001218; Sulfotransferase; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 295 AA; 34174 MW; A0898231C36B06C6 CRC64;  
 CC  
 Query Match 28.0%; Score 429; DB 2; Length 295;  
 Best Local Similarity 33.2%; Pred. No. 2.2e-27;  
 Matches 95; Conservative 57; Mismatches 114; Indels 20; Gaps 4;  
 QY 17 SKVEFHGVLPPFCRGKMEETANPPRSDVWIVTPKSGTSLLOEVVYLVSGADPDE 76  
 Db 12 AKYVK--GIPLIKYFAELGPLESEFEAWPDDLLISTYPKSGTWTWSEILDLIYQSGDLEK 69  
 QY 77 IGLMIDEQLPVLEYPQP----GLDIIKELTSPRLIKSHLPYRFLPSPDLHNGDSKVIYMA 132  
 Db 70 CORAPVFLRVPLEFSAPGVPTGVLLKDTTPAPRLKTHLPALLKPTLLDQKVKVYIA 129  
 QY 133 RNPDLVSVYQFHRSLRTMSYRGTFQFCRFPMNDKLGYSWFEHVFQEFWEHRMDSNVL 192  
 Db 130 RNAKDVAVSYHYFHYRMAKVPDPGTWDSPLFXFMAGVCYGSYQHVQEWELSHTHPVL 189  
 QY 193 FLKYEDMRDLVTWQLARFLGVSCDKAQLAAL---TEHCQLVDQCCNAALPVG--- 246  
 Db 190 YLFYEDIKEDSKREIQKLEFGRSLPEETVHVIVQRTSFKEMKKNPMTNYSTIPTAVMD 249  
 QY 247 -----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTDFDYL 284  
 Db 250 HSIASFMRKGITGDWKSFTVAQNELFEAHYAKWAGCKLRFERWEL 295  
 RESULT 62  
 Q90WR6 CHICK PRELIMINARY; PRT; 307 AA.  
 AC Q90WR6  
 DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.  
 DT 07-FEB-2006, sequence version 1.  
 DE Sulfotransferase 1C.  
 GN Name=SULT1C;  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Liver;  
 RX PubMed=15234270; DOI=10.1016/j.abb.2004.05.008;  
 RA Wilson L.A., Reynolds G.E., Darras V.M., Coughtrie M.W.H.;  
 RT "cDNA cloning, functional expression, and characterization of chicken  
 RT sulfotransferases belonging to the SULT1B and SULT1C families.";  
 RL Arch. Biochem. Biophys. 428:64-72(2004).  
 CC  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC  
 CC EMBL: AJ416889; CAC95180.1; -; mRNA.  
 DR HSSP; P50224; ICDM.  
 DR Ensembl; ENSGALG00000016805; Gallus gallus.

DR GO: GO:0008146; F:sulfotransferase activity; IEA.  
 DR GO: GO:0016740; F:transferase activity; IEA.  
 DR InterPro: IPR000863; Sulfotransferase.  
 DR Pfam: PF00685; Sulfotransferase\_1.  
 DR ProDom: PD001218; Sulfotransferase; 1.  
 KW Transferase.  
 SQ SEQUENCE 307 AA; 36251 MW; 6B76EB00EDA04C9C CRC64;  
 CC  
 Query Match 28.0%; Score 428; DB 2; Length 307;  
 Best Local Similarity 32.6%; Pred. No. 2.8e-27;  
 Matches 87; Conservative 61; Mismatches 97; Indels 22; Gaps 2;  
 QY 36 BEIANFPVRPSDVWIVTPKSGTSLLOEVVYLVSGADPDEIGLMNIDEQLPVLEYPQP- 94  
 Db 37 DQVWNFKARPDLLVATYAKAGTTWTQIVDMIQQNGDIEKRRASTYKRHPFLEWYIPD 96  
 QY 95 -----GLDIIKELTSPRLIKSHLPYRFLPSPDLHNGDSKVIYMARNPDLVSVYQFHR 147  
 Db 97 SSPGLYSGLKLAEMPSPTWKTHLPVLVPPSPFWEQNKIIYVARNAKDNLVSYHFHR 156  
 QY 148 SLRTMSYRGTFQFCRFPMNDKLGYSWFEHVFQEFWEHRMDSNVLFLEKYMHRDLVTMV 207  
 Db 157 MNKVLDPDGTTEETFEKPMNGEVLWGSWYDHVKGWAKAKHRIILYLFVDMKENPKREI 216  
 QY 208 EQLARFLGVSCDKAQL-----EALTEHCQLVDQCCNAALPVGGRGVGLW 253  
 Db 217 QKIMKLEKDLDEEVLNKKIINYTSFEMKDNPMNTNYTKDFVGMVDHVSVPFMRKGSVGDW 276  
 QY 254 KDIFTVSMNEKFDLVYKQKMGKCDLTTF 280  
 Db 277 KNYFTVALNKFPDQDYKKGMADTSLVF 303  
 RESULT 63  
 STLS3 BRARE  
 ID STLS3 BRARE STANDARD; PRT; 301 AA.  
 AC Q772V2;  
 DT 31-AUG-2004, integrated into UniProtKB/Swiss-Prot.  
 DT 01-OCT-2003, sequence version 1.  
 DE 07-FEB-2006, entry version 19.  
 DE Cytosolic sulfotransferase 3 (EC 2.8.2.-) (SULT1 ST3).  
 GN Name=sult1st3;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [MRNA], ACTIVITY, TISSUE SPECIFICITY, AND ENZYME  
 RP REGULATION.  
 RX MEDLINE=22630056; PubMed=12745256; DOI=10.1016/S0003-9861(03)00172-3;  
 RA Sugahara T., Liu C.-C., Carter G., Pai T.G., Liu M.-C.;  
 RT "cDNA cloning, expression, and functional characterization of a  
 RT zebrafish SULT1 cytosolic sulfotransferase.";  
 RL Arch. Biochem. Biophys. 414:67-73(2003).  
 CC  
 CC -!- FUNCTION: Catalyzes the conjugation of sulfate to a variety of  
 CC xenobiotic and endogenous compounds, including dopamine, T3  
 CC (triiodo-L-thyronine), T4 (thyroxine), estrone, DHEA  
 CC (dehydroepiandrosterone), flavonoids, isoflavonoids and other  
 CC phenolic compounds.  
 CC -!- ENZYME REGULATION: Inhibited by Hg(2+), Co(2+), Zn(2+), Cd(2+),  
 CC Cu(2+) and Pb(2+) ions. Activated slightly by Mn(2+), Ca(2+) and  
 CC Mg(2+) ions.  
 CC -!- BIOPHYSICO-CHEMICAL PROPERTIES:  
 CC pH dependence:  
 CC Optimum pH is about 5.0 with n-propyl gallate as substrate, and  
 CC another smaller pH optimum is observed spanning pH 9.5-10.5.  
 CC Optimum pH is 10.5 with dopamine as substrate;  
 CC Temperature dependence:  
 CC Active from 20 to 43 degrees Celsius;  
 CC -!- SUBCELLULAR LOCATION: Cytoplasm.  
 CC -!- SIMILARITY: Belongs to the sulfotransferase family.  
 CC -----



```
RESULT 65
Q4F2P1_XENLA PRELIMINARY; PRT; 302 AA.
ID Q4F2P1;
AC Q4F2P1;
DT 30-AUG-2005, integrated into UniProtKB/TrEMBL.
DT 30-AUG-2005, sequence version 1.
DT 07-FEB-2006, entry version 2.
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RC TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.,
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altchul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RA Klein S., Gethard D.S.;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC099307; AAH99307.1; -; mRNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Hypothetical protein.
SQ SEQUENCE 302 AA; 35792 MW; AF91BD3FB8BF90C9 CRC64;

Query Match 27.9%; Score 427.5; DB 2; Length 302;
Best Local Similarity 31.6%; Pred. No. 3e-27;
Matches 89; Conservative 61; Mismatches 115; Indels 17; Gaps 3;

QY 18 KYFEPHGVRLPPFCRGKMEIANFPVPSPDVWIVTPKSGTSLLOEVVLYVLSQCADPDEI 77
ID Q6IM16_HUMAN PRELIMINARY; PRT; 304 AA.
DB 4 KLOSMGIRIAGVIATNNQIIRTFQARGLDVLVIATPKSGTWTWQEIIVDLINEGNEEIC 63
QY 78 GLMNIDEQLP---VLEYQPQGLDIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIMARN 134

RESULT 66
Q2XV97_MACEA PRELIMINARY; PRT; 294 AA.
ID Q2XV97_MACEA PRELIMINARY; PRT; 294 AA.
AC Q2XV97;
DT 20-DEC-2005, integrated into UniProtKB/TrEMBL.
DT 20-DEC-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Sulfotransferase family 1E, estrogen-preferring, member 1.
GN Name=SUH1E1;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopitheidae; Cercopitheciae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Liu H., Larbie F., Luu-The V.;
RL Submitted (SEP-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; DQ228171; ABB76812.1; -; mRNA.
DR GO; GO:0016740; F:transferase activity; IEA.
KW Transferase.
SQ SEQUENCE 294 AA; 35159 MW; B9997AFD8CC353C0 CRC64;

Query Match 27.6%; Score 423; DB 2; Length 294;
Best Local Similarity 31.9%; Pred. No. 6.9e-27;
Matches 91; Conservative 58; Mismatches 118; Indels 18; Gaps 3;

QY 14 EFESKYFEPHGVRLPPFCRGKMEIANFPVPSPDVWIVTPKSGTSLLOEVVLYVLSQCAD 73
ID Q6IM16_HUMAN PRELIMINARY; PRT; 304 AA.
DB 6 DYENFELHGVLMYKDFVYKYNVDVETTFQARPDVLVIATPKSGTWTWSEIAYMIYKGD 65
QY 74 PDEIGLMNIDQLPVLVYPPQ---GLDIKELTSPRLIKSHLPYRFLPSDLHNGDSKVI 129
DB 66 VEKCKEDVFNRIFFLECKEDLMNGVKQLDEMSPRIKTHLPPELLPASFEWKNCKII 125
QY 130 YMARNPDLVVSYYQFHRSLTMSYRGTFQFCRRFMNDKLGYSWFEHVFQEFWEHRMDS 189
DB 126 YLCRNAKDVAVSFYVFLWAGHPGSPFEVKEFMQGVPGYSGWYKHKVSWWEKESP 185
QY 190 NVLFKVEDMHRDLVTWVEQLARFLGVSCDQAQLEALTEHCH-----QLVQD 236
DB 186 RILFLFYEDLKEDIRKEVIKLIHFLERKPSBELVDKIHTHSFQEMKNPNSTNYTTLPE 245
QY 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 246 IMNOKVSFFMRKKGITGDWKHFTVALNEKFDKHFEQMKESTLKF 290

RESULT 67
Q6IM16_HUMAN PRELIMINARY; PRT; 304 AA.
ID Q6IM16_HUMAN PRELIMINARY; PRT; 304 AA.
AC Q6IM16;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
```

```
DT 05-JUL-2004, sequence version 1.
DE 07-FEB-2006, entry version 10.
DE SULTIC3 splice variant d.
GN Name=SULTIC3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14676822; DOI=10.1038/gt.tj.6500223;
RA Freimuth R.R., Weipert M., Chute C.G., Wieben E.D., Weinshilboum R.M.;
RT "Human cytosolic sulfotransferase database mining: identification of
RT seven novel genes and pseudogenes.";
RL Pharmacogenomics J. 4:54-65(2004).
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; BK001432; DAA01771.1; -; Genomic DNA.
CC HSP; P49891; 1A0V.
CC Ensembl; ENSG00000196228; Homo sapiens.
CC GO; GO:0008146; P:sulfotransferase activity; IEA.
CC InterPro; IPR000863; Sulfotransferase.
CC Pfam; PF00685; Sulfotransferase_1; 1.
CC ProDom; PD001218; Sulfotransferase; 1.
CC SEQUENCE 304 AA; 35889 MW; 5510C869578BB70C CRC64;

Query Match 27.6%; Score 422; DB 2; Length 304;
Best Local Similarity 32.3%; Pred. No. 8.7e-27;
Matches 97; Conservative 59; Mismatches 124; Indels 20; Gaps 5;

QY 1 MAESAETPTSGEPE-SKYEFHGVRLPPFCRCGMEEIANPPVPSPDVWIVTPKSGTS 59
DB 1 MAKIEKNAFTWEKKELFNEMVDGPTLLSKWEKVCNFOAPKFDLLILATYPSGTT 60

QY 60 LQEVVYLVSQGADPEIGLNMIDQLPVLE--YP---OPGLDIKEITSPRLIKSHLPY 114
DB 61 WHHEILMDLNDGDVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQILKTHLPS 120

QY 115 RPLPSDLHNGDSKVLYMARNKDLVYVYQFHRSLRTMSYRGTFOEFCRRFMNDKLGYS 174
DB 121 HLIPPSIWKENCKIYIVARNPKDCLVSYVYHFRHMASFMPDPPQNLSEFYKFGKSVGVGS 180

QY 175 WPEHVQEFWEHRMDSNVLPFLKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCHQLV 234
DB 181 WFDHVKGWAAKDMHRILYLFYEDIKKOPKRIEIKLFLKXDISSEILNKIIYHTSFDV 240

QY 235 ---DQCNAEALPVG-----RGVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
DB 241 MKQNPMTNVTTLPTSIMDHSISPFMRKMGPGDKNVFTVAQNEEFKDYQKMGAGSTLTF 300

RESULT 68
Q6XZC1 BRARE PRELIMINARY; PRT; 304 AA.
AC Q6XZC1
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DE 07-FEB-2006, entry version 15.
DE SULT1 sulfotransferase isoform 4.
GN Name=sult1st4; ORFNames=zgc:76922.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Liu M.-C., Liu C.-C., Sugahara T.;
```

```
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richardson D., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield V.S.N., Krywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RG NIH MGC Project;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AY196986; AAP55638.1; -; mRNA.
CC EMBL; BC066584; AAH66584.1; -; mRNA.
CC HSP; P49891; 1A0V.
CC Ensembl; ENSDARG00000003181; Danio rerio.
CC ZFIN; ZDB-GENE-040426-2054; sult1st4.
CC GO; GO:0008146; P:sulfotransferase activity; IEA.
CC GO; GO:0016740; P:transferase activity; IEA.
CC InterPro; IPR000863; Sulfotransferase.
CC Pfam; PF00685; Sulfotransferase_1; 1.
CC ProDom; PD001218; Sulfotransferase; 1.
CC KW Transferase.
CC SEQUENCE 304 AA; 35556 MW; 20BE69D9DC409873 CRC64;

Query Match 27.5%; Score 421.5; DB 2; Length 304;
Best Local Similarity 34.9%; Pred. No. 9.6e-27;
Matches 98; Conservative 51; Mismatches 111; Indels 21; Gaps 6;

QY 20 FEFHGVRLPPFCRCGMEEIANPPVPSPDVWIVTPKSGTSLQEVVYLVSQGADPEIGL 79
DB 21 FDFEGVLTREFTDNWENVKNFOARPPDDILATYPKAGTTWVSYLDLLYFGSDENQTS- 79

QY 80 MNIDEQLPVLE-----YQPGLDIIKEL-TSPRLIKSHLPVRLPLPSDLHNGDSKVYMA 132
DB 80 QPIVQVRFLESCFOEFSTIGTEMADNLPTSPRLIKTHLPVLQVPKSFWEQNSRVVYA 139

QY 133 RNPKDLVSVYQFHRSLRTMSYRGTFOEFCRRFMNDKLGYSWFEHVQEFWE-HRMSDNV 191
DB 140 RNAKDNAVSYFHFDRMNMVQDPDGDWSDYLDKFMQGVNFGVSWFDHVGWQKRSYPNM 199

QY 192 LFLKYEDMHRDLVTWVEQLARFLGVS CDKAQLEALTEHCH---QLVDQCNAEALP---- 244
DB 200 LYMPFEDISEDTGREVNRLCPSFLGLSTSVQKEKITKGQVQFDAMKQNTLINHTVTPFLDC 259

QY 245 -----VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
DB 260 KISPFMRKGVGDWKSHTVAQNERFDEVYKQMKNSGVTF 300

RESULT 69
```



OC Muroidea; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
[1]  
RN NUCLEOTIDE SEQUENCE [MRNA], AND PROTEIN SEQUENCE OF 10-20 AND 28-38.  
RP STRAIN=Sprague-Dawley; TISSUE=Liver;  
RX MEDLINE=94043328; PubMed=8227031;  
RA Nagata K., Ozawa S., Miyata M., Shimada M., Gong D.-W., Yamazoe Y.,  
RA Kato R.;  
RT "Isolation and expression of a cDNA encoding a male-specific rat  
RT sulfotransferase that catalyzes activation of N-hydroxy-2-  
RT acetylaminofluorene.";  
RL J. Biol. Chem. 268:24720-24725(1993).  
[2]  
RN NUCLEOTIDE SEQUENCE [MRNA].  
RX MEDLINE=94306556; PubMed=8033246; DOI=10.1016/0009-2797(94)90057-4;  
RA Yamazoe Y., Nagata K., Ozawa S., Kato R.;  
RT "Structural similarity and diversity of sulfotransferases.";  
RL Chem. Biol. Interact. 92:107-117(1994).  
CC -!- FUNCTION: Sulfation of phenols and bioactivation of N-  
CC hydroxyarylamines. Is responsible for the formation of N-hydroxy-  
CC 2-acetylaminofluorene, a reactive metabolite which exhibits  
CC toxicity by binding to DNA, RNA and protein.  
CC -!- SUBCELLULAR LOCATION: Cytoplasm (potential).  
CC -!- TISSUE SPECIFICITY: Liver. Male >> Female.  
CC -!- DEVELOPMENTAL STAGE: Male specific. Maximum at 9 weeks and  
CC maintained in 9-month-old rats. Can be detected at low level in  
CC females up to 9-week-old rats but then decreases to undetectable  
CC level.  
CC -!- INDUCTION: Induced by estrogens and suppressed by androgens.  
CC -!- SIMILARITY: Belongs to the sulfotransferase family.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL: L22339; AAA42181.1; -; mRNA.  
DR PIR: A49098; A49098.  
DR HSSP: P50224; 1CJM.  
DR Ensembl: ENSRNOG00000011269; Rattus norvegicus.  
DR LinkHub: P50337; -;  
DR GO: GO:0005829; C:cytosol; NAS.  
DR GO: GO:004062; F:aryl sulfotransferase activity; IDA.  
DR InterPro: IPR000863; Sulfotransferase.  
DR Pfam: PF00685; Sulfotransfer\_1.  
DR ProDom: PD001218; Sulfotransferase; 1.  
KW Direct protein sequencing; Transferase.  
FT CHAIN 1 304 N-hydroxyarylamines sulfotransferase.  
FT /FTID=PRO 0000085138.  
FT NP\_BIND 56 61 PAPS (By similarity).  
FT NP\_BIND 139 147 PAPS (By similarity).  
FT NP\_BIND 202 238 PAPS (By similarity).  
FT NP\_BIND 266 268 PAPS (By similarity).  
FT ACT\_SITE 117 117 Proton acceptor (By similarity).  
SQ SEQUENCE 304 AA; 35764 MW; C772B2EA7BD74198 CRC64;  
Query Match 27.4%; Score 419.5; DB 1; Length 304;  
Best Local Similarity 32.6%; Pred. No. 1.4e-26;  
Matches 91; Conservative 57; Mismatches 11; Indels 19; Gaps 4;  
QY 21 EFHGVRLPPFCRGKMEIANFPVPSDVWIVTYPKSGTSLQEVVYLVVSQADPDEIGLM 80  
DB 22 EVNGILMSKLSMDNDWKIWNFOAKPDDLIIATYAKAGTTWTQEIIVDMIQNDGVDQKQORA 81  
QY 81 NIDQLPVLVEY--POP-----GLDIKELTSPLIKSHLPYRFLPSDLNGLNDSKVIYMARNP 135  
DB 82 NYDRHPPIETWLPSPNLNSGLDLANKMPSPTLTKTHLPVHMLPPSPFWKENSKIIVARNA 141  
QY 136 KDLVSVYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVEHWRMDSNVLFK 195  
DB 142 KDLVSVYYFRRMKNMLPDGTLGIEYIEQFAGKVLGWSWDVHVGWWDVQDQHRILYLF 201  
QY 196 YEDMRDLVTWVEQLARFLGVSCDKAQLEALTEHCHQLV---DQCNEALP-----244  
DB 202 YEDNKEDPKREIKKIAKFLKEDISEVLNKKIYHTSFDVMKENPMANVTTLTLPSSIMDSHI 261

QY 245 ---VGRGRVGLWKDIFTVMNEKEFDLVYKQMGKCDLTF 280  
DB 262 SPFNRKMGPGDKNKFYTFVAQSEDFEDYRRKMAGSNITF 300

## RESULT 72

Q5M8B5 RAT  
ID Q5M8B5 RAT PRELIMINARY; PRT; 304 AA.  
AC Q5M8B5;  
DT 01-FEB-2005, integrated into UniProtKB/TrEMBL.  
DT 01-FEB-2005, sequence version 1.  
DT 07-FEB-2006, entry version 8.  
DE Sultic1 protein.  
GN Name=Sultic1;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muroidea; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN (1)  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Liver;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.B.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Liver;  
RG NIH MGC Project;  
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL: BC088125; AAH88125.1; -; mRNA.  
DR GO: GO:0008146; F:sulfotransferase activity; IEA.  
DR InterPro: IPR000863; Sulfotransferase.  
DR Pfam: PF00685; Sulfotransfer\_1.  
DR ProDom: PD001218; Sulfotransferase; 1.  
SQ SEQUENCE 304 AA; 35774 MW; 961E7D24E9B8BFOA CRC64;

Query Match 27.4%; Score 419.5; DB 2; Length 304;  
Best Local Similarity 32.3%; Pred. No. 1.4e-26;  
Matches 90; Conservative 57; Mismatches 113; Indels 19; Gaps 3;  
QY 21 EFHGVRLPPFCRGKMEIANFPVPSDVWIVTYPKSGTSLQEVVYLVVSQADPDEIGLM 80  
DB 22 EVNGILMSKLSMDNDWKIWNFOAKPDDLIIATYAKAGTTWTQEIIVDMIQNDGVDQKQORA 81  
QY 81 NIDQLPVLVEYPOP-----GLDIKELTSPLIKSHLPYRFLPSDLNGLNDSKVIYMARNP 135  
DB 82 NYDRHPPIETWLPSPNLNSGLDLANKMPSPTLTKTHLPVHMLPPSPFWKENSKIIVARNA 141  
QY 136 KDLVSVYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVEHWRMDSNVLFK 195

```
Db 142 KDLVSYYYFSRNKMLPPDGLTGEYIEQFKAGKVLGWSYDHWKGMWDVKDQHRILYLF 201
Qy 196 YEDMRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLV---DQCNAEALP----- 244
Db 202 YEDMKEDPKREIKKAKTAKFLEKDISSEVLNKKIYHTSFVWKENPMANYTTLSSIMDHSI 261
Qy 245 ----VGRGVLWKDIFTVSMNKKFDLVYKQKMGKCDLTF 280
Db 262 SPFMKMGPGDKWNTFTVAQSEDFDEYRRKMAGSNITF 300

RESULT 73
Q68EV4_XENTR
ID Q68EV4_XENTR PRELIMINARY; PRT; 287 AA.
AC Q68EV4;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 14.
DE Hypothetical protein MGC75696.
GN Name=MGC75696;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Ustin T.B., Toshlyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.; to the EMBL/GenBank/DBSJ databases.
RL Submitted (NOV-2003) to the EMBL/GenBank/DBSJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC061263; AAH61263.1; -; mRNA.
DR HSSP; F49891; IAQU.
DR Ensembl; ENSXETG000000012821; Xenopus tropicalis.
DR GO; GO:0008146; F:sulfofransferase activity; IEA.
DR InterPro; IPR000863; Sulfofransferase.
DR Pfam; PF00685; Sulfofransfer 1; 1.
DR ProDom; PD001118; Sulfofransferase; 1.
DR Hypothetical protein.
SQ SEQUENCE 287 AA; 33145 MW; F5F382031752B4A8 CRC64;

Query Match 27.2%; Score 416.5; DB 2; Length 287;
Best Local Similarity 34.6%; Pred. No. 2.3e-26;
Matches 94; Conservative 48; Mismatches 111; Indels 19; Gaps 4;
```

```
Qy 27 LPPFCRGKMBEIANFFVPRPSPVWIVTYPKSGTSLLOEVVYLVSQGADPEITGLMMDIROL 86
Db 15 LGPEA-ANWENVKFOARPDLLIATYPKSGTWMSEIVDQIVAVSNSEKCTAAIYBRV 73
Qy 87 PVLEYPOP---GLDIIKELTSPRLIKSHLPVRELPSPDLHNGDSKVIYMARNPDKLVSY 142
Db 74 PFLEYAVPDMSPGTOALDQRASPRLIKTHLPVELLPKSFWDNKVKVIIVARNKQDVAVSY 133
Qy 143 YQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFVEHVFWEHRMDSNVLFKLYEDMHRD 202
Db 134 YHFYRMAIVHPPEPGTWDFLDSYINGKVCFGSWSAHVKGWQKAKEDVDLYLFYEDMLED 193
Qy 203 LVTWVEQLARFLGVSCDKAQLAEAL---TEHCHQLVDQCNAEALP-----VGRG 248
Db 194 PTREIRKVKVFMGKDLPEETVEKIASQTSFKAMQKNLSNYSMVPSVMDHISIPFMRKG 253
Qy 249 RVGLWKDIFTVSMNKKFDLVYKQKMGKCDLTF 280
Db 254 VGDWKNQFTVAQNEKFDDEYQREMSDGLSIF 285

RESULT 74
Q68EV4_XENLA
ID Q68EV4_XENLA PRELIMINARY; PRT; 276 AA.
AC Q68EV4;
DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE MGC84291 protein.
GN Name=MGC84291;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Ustin T.B., Toshlyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Klein S., Gerhard D.S.; to the EMBL/GenBank/DBSJ databases.
RL Submitted (AUG-2004) to the EMBL/GenBank/DBSJ databases.
```





```

Db 150 PEPGTWEELEKPMAGQVSGPWYDHVKGWWEKREYRILYLFYEDMKENPKCEIQOILK 209
QY 213 FLGVSCDQKQLEALTEHCH-----QLVDQCCNABALPVGGRVGLWKDIFT 258
Db 210 YLEKDIPEELINKILYHSSVNMKENPSANYITWKEEMDHVSPPMRKGISGDWKNQFT 269
QY 259 VSMNEKFDLVYKQKMGKCDLTF 280
Db 270 VAQYKFEEDYVKKMEDSTLKF 291

RESULT 77
QSEAWO XENLA
ID QSEAWO XENLA PRELIMINARY; PRT; 303 AA.
AC QSEAWO;
DT 15-MAR-2005, integrated into UniProtKB/TrEMBL.
DT 15-MAR-2005, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Egg;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Egg;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Egg;
RA Klein S., Gerhard D.S.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; BC090224; AAH90224.1; -; mRNA.
DR GO; GO:0008146; F:sulfoltransferase activity; IEA.
DR InterPro; IPR000863; Sulfoltransferase.
DR Pfam; PF00685; Sulfoltransferase 1; 1.
DR ProDom; PD001218; Sulfoltransferase; 1.
DR Hypothetical protein.
SQ SEQUENCE 303 AA; 35647 MW; B940559678751DA1 CRC64;

```

```

Query Match 27.0%; Score 413; DB 2; Length 303;
Best Local Similarity 32.1%; Pred. No. 4.9e-26;
Matches 89; Conservative 55; Mismatches 115; Indels 18; Gaps 3;
QY 22 FHGVRLPPFCRCMBEIANFPVRPSDWJVTVPKSGTSLQLQVVLVSGADPDEIGLWN 81
Db 23 FRGVPMAKMNVENWEPVEYFQARHDDVVLATYPKAGTTWVSEILDIMYNGDLEKQORD 82
QY 82 IDEQLPVLEYPOPGL----DIIKELTSRLIKSHLPYRPLPSDLHLNGDSKVIYMARNP 137
Db 83 IYNRPVMEIRIPGMPGVDQLELLASPLRIKTHLPQLMPSEFWEKKCKVIYVARNAK 142
QY 138 LVVSYVYQFHRSLRSTSYRGTFQFCRRFMNDKLGYSWFEHVQEFWEHRMDSNVLFKYE 197
Db 143 VAVSYFFHFQMVKALPDGPDWKLADYMGVTSYGSWYDHVKGWWEKREKQYGLYLFYE 202
QY 198 DMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCHQIV---DOCCNABALP----- 244
Db 203 DLKEDPKREIKKILHYLEKLSDEVLEKIVVHHTSFQVMSKNDWMANYKTIPNDILNQNTA 262
QY 245 -VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 263 FMRKGEAGDWKNHFTVAQNEAPDNQYQSQMLGTSLHF 299

```

```

RESULT 78
Q80VR3_MOUSE
ID Q80VR3_MOUSE PRELIMINARY; PRT; 304 AA.
AC Q80VR3;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 18.
DE Sulfoltransferase family, cytosolic, 1C, member 1.
GN Name=Sult1c1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Olfactory epithelium;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Olfactory epithelium;
RA Director MGC Project;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----

```

Query Match 26.9%; Score 412; DB 1; Length 294;  
Best Local Similarity 33.1%; Pred. No. 5.7e-36;  
Matches 91; Conservative 56; Mismatches 108; Indels 20; Gaps

Db 12 AKYVK--GIPLIKYFAEALGPLSEAFNPDDLLISTIPKSTGTWVSEILDIIYQEGDLEK 60  
Qy 77 IGLMNIDEQLPVLEYQPQ-----GLDI IKELTSPLRIKSHLPIRYFLPSDLNGDSKVITYMA 130  
Db 70 CORAPVFLVPFLEFSAGPVTGPTEGVELLKDTTPAPRLTKTHLPALLPKTLDPQKVVIYA 122  
Qy 133 RNPXDLVSVYYOYFHRSIRMTSYRGTFQBFRCRRFMNDKLGYGSWPFHVOEFWEHRMDSNVL 198  
Db 130 RNAKDVAVSYYIFHYRMAKHVHPDGPTWDSELFKEFPAGEVCYCSWYOHVQEWELSHTHPVL 180  
Qy 193 FLKYEDMHDRDLVTWVEYQIARFLGVGSCDKAQLAL---TEHCHQLVDQCCEAAELPVG--- 240  
Db 190 YLFYEDIKEDPKREIQILEFTIGRSLPBETVDHI VQRTSFKEMKNPMNTYSTIPTAYMD 242  
Qy 247 -----RGRVGLMKIDIFTYSMNEKFDDLIVYKKQM 273  
Db 250 HSIAPMRKGITGDWKSKTFTTAQNELFEAHYAKKM 284

RESULT 80  
Q7S293\_XENLA  
ID ID Q7S293\_XENLA PRELIMINARY; PRT; 294 AA.  
AC Q7S293\_ AC  
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.  
DT 01-OCT-2003, sequence version 1.  
DT 07-FEB-2006, entry version 14.  
DE MGC64389 protein.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
OC Xenopodinae; Xenopus; Xenopus.  
NCBI\_Taxid=8355;  
RX NCBI [1]  
RN NUCLEOTIDE SEQUENCE.  
RP TISSUE=Whole;



OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Mus.  
 OK NCBI\_TaxID=10090;  
 [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Methods Enzymol. 303:19-44(1999).  
 [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;  
 RX PubMed=16141072; DOI=10.1126/science.1112014;  
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,  
 RA Oyama R., Ravasi B., Lenhard B., Wells C., Kodaira K., Tomizawa K.,  
 RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,  
 RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,  
 RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,  
 RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,  
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,  
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,  
 RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,  
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gabibidi M.,  
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,  
 RA Guscinich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,  
 RA Hill D., Hummel M., Iacono M., Ikeo K., Iwama A., Ishikawa T.,  
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,  
 RA Kitano H., Kollis G., Krishnan S.P., Kruger A., Kummerfeld S.K.,  
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,  
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,  
 RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,  
 RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,  
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,  
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,  
 RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,  
 RA Roest B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,  
 RA Schonbach C., Sekiguchi K., Sempke C.A., Seno S., Sessa L., Sheng Y.,  
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,  
 RA Sperling S., Stupka E., Sugtara K., Sultana R., Takeda Y., Taki K.,  
 RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,  
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,  
 RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,  
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,  
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,  
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,  
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,  
 RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,  
 RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,  
 RA Tagami M., Waki K., Watanabe A., Okamura-Oho Y., Suzuki H., Kawai J.,  
 RA Hayashizaki Y.;  
 RT "The transcriptional landscape of the mammalian genome.";  
 RL Science 309:1559-1563(2005).  
 [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;  
 RX PubMed=16141073; DOI=10.1126/science.1112009;  
 RG RIKEN Genome Exploration Research Group, and Genome Science Group  
 (Genome Network Core Team) and the FANTOM Consortium;  
 RT "Antisense Transcription in the Mammalian Transcriptome.";  
 RL Science 309:1564-1566(2005).  
 [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,

GAasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gutincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lennard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Pavlovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Petroski T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kigawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RT Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Iwama M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 403:685-690(2001).  
 [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;  
 RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 [7]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaehiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 [8]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;  
 RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,

RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fahney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA	Butterfield Y.S.N., Kryzyski M.I., Skalska U., Smalhus D.E.,
RA	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences.";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
EN	[2]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=C57BL/6J; TISSUB=Kidney;
RA	Director MGC Project;
RL	Submitted (FEB-2004) to the ENBL/GenBank/DBDJ databases.
CC	-----
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC	Distributed under the Creative Commons Attribution-NoDerivs License
CC	-----
CC	ENBL; BC066190; AAH66190.1; -; mRNA.
DR	HSP; P49891; 1AQU.
DR	Ensembl; ENSMUSG0000029273; Mus musculus.
DR	MGI; MGI:1926341; Sultld1
DR	GO; GO:0004062; F:aryl sulfotransferase activity; IDA.
DR	GO; GO:0000103; P:sulfate assimilation; IDA.
DR	InterPro; IPR000863; Sulfotransferase.
DR	Pfam; PF00685; Sulfotransfer_1; 1.
DR	ProDom; PD001218; Sulfotransferase; 1.
DR	Transferase.
KW	SEQUENCE 295 AA; 35011 MW; A569FAF60CC0716F CRC64;
QY	Query Match 26.9%; Score 412; DB 2; Length 295;
QB	Best Local Similarity 31.7%; Pred. No. 5.7e-26;
MB	Matches 83; Conservative 63; Mismatches 98; Indels 18; Gaps 2
Qy	37 EIANFPVRSDVIVTVPKSGTSLQEVIYVSOGADPDEGLMNIDQLPVLEYPPQGL 96
Db	:: ::
Qy	30 QVESFEARPDILISTYPKSGTTWVSEILDLYNNGDAEKCCKDAIKRVPFMEIIPGI 89
Db	:: ::
Qy	97 ----DIIEKLTSPRIKSHLPYRFLPSDLHGDSKVIMARNPKDVLVSYQFHRSLTRM 152
Db	:: ::
Qy	90 TNGVMNNMPSPRIVATHLPVQLPSSFWNKDCKIIVARNAKDVVVSYFYQMAMKH 149
Db	:: ::
Qy	153 SYRGTFQFCRRFMNDKLGISGFHEHQEFWEHRMDSNVLFKYEDMHRDLVTMVQELAR 212
Db	:: ::
Qy	150 PEPTGWEEFLKFAGQVSGPWYDHVKSWEKEKEYRILYLFYEDMKENPCEIQKILK 209
Db	:: ::
Qy	213 FLGVSCDKQAQLAEALTEHC-----QLVDCCNAEALPVGRGRVGLWKDIFT 258
Db	:: ::
Qy	210 FLEKDIPETILTINKLIYHSFSVFMKGNPANYTTWMKEBMDHSVPFMRKGISGDWNQFT 269
Db	:: ::
Qy	259 VSMNEKFDLVYKMGKCDLTF 280
Db	:: ::
Qy	270 VAQYEKFEEDYVKKMEDSTLK 291
Db	:: ::
RESULT 84	
Q9ZIG0_RAT	
ID	Q9ZIG0_RAT PRELIMINARY; PRT; 308 AA.
AC	Q9ZIG0;
DT	01-MAY-1999, integrated into UniProtKB/TREMBL.
DT	07-FEB-2006, entry version 1.
DE	Tyrosine-ester sulfotransferase (EC 2.8.2.9).
OS	Name=Sultld1;
GN	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC	Muridea; Murinae; Rattus.
OX	NCBI_TaxID=10116;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=Sprague-Dawley; TISSUE=Kidney;

```

RA Herrmann A., Stoffel W.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBDJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U32372; AAC99890.1; -; mRNA.
DR HSSP; P49888; 1HY3.
DR Ensembl; ENSRNOG0000001960; Rattus norvegicus.
DR RGD; 620491; Sult1d1.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0017067; F:tyrosine-ester sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR Transferrase.
SQ SEQUENCE 308 AA; 36341 MW; 68719C92233EE9AE CRC64;

Query Match      26.9%; Score 412; DB 2; Length 308;
Best Local Similarity 32.3%; Pred. No. 6.1e-26;
Matches 93; Conservative 59; Mismatches 110; Indels 26; Gaps 5;

QY 15 PESKFEFHGRLPFCRGMKEIEANFVRPSDVWIVTPKSGTSLLOEVVYLVNSQGDGP 74
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 8 FRRELVDVQGIPLFWSIAEQNSQVESFEARPDILISTYPKSGTTWISILDLYNNGDA 67
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 75 DEIGLMNIDQLPVLEYQPGSL-----DIKELTSPRLKSHLPYRFLPSDLHNGDSKVIY 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 68 EKCKRDAIYRRVFPFMELIIPGITNGVEMLDNMQSPRLVKTHLPVQLPSPSFWNRDCKMIY 127
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 131 MARNPDKLVVSYQFHRSLRTMSYGTQFCRCRRFMNDKLGYSWFHVFQFWEHRMDSN 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 128 VARNAKDVAVSYFYHQWAKHPEPTWEEFLEKFMAGVSPGPDYDHVKGWKEKRYR 187
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 191 VLFLKYEDMHRDLVTWVQLARFL-----GVSCDKAQLA-LTEHCHQL 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 168 ILICYEDMKEDPKCEIQKVLKFLKDIPEVVNVKILYHSSFSVMKNPANSYTTMMKEE 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 234 VDQCNAALP-VGRGRVGLWKDITVSMNEKFDLVYKQMGKCDLTF 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 MDQSVS----PPMRKGISGDMKNQFTVAQYEKFEEDYVKMGKESTLKF 291
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 85
Q9R2C2 MOUSE
ID Q9R2C2_MOUSE PRELIMINARY; PRT; 309 AA.
AC Q9R2C2;
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Tyrosine-ester sulfotransferase (EC 2.8.2.9).
GN Name=Sult1d1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BALB/c; TISSUE=Kidney;
RA Herrmann A., Stoffel W.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBDJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U32371; AAC99889.1; -; mRNA.
DR HSSP; P49888; 1HY3.
DR Ensembl; ENSMUSG000000029273; Mus musculus.
DR MGI; MGI:1926341; Sult1d1.
DR GO; GO:0004062; F:aryl sulfotransferase activity; IEA.
DR GO; GO:0000103; P:sulfate assimilation; IEA.
DR InterPro; IPR000863; Sulfotransferase.

```

```

DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR Transferrase.
SQ SEQUENCE 309 AA; 36713 MW; E5D5E1CD3C6D5117 CRC64;

Query Match      26.9%; Score 412; DB 2; Length 309;
Best Local Similarity 31.7%; Pred. No. 6.1e-26;
Matches 83; Conservative 63; Mismatches 98; Indels 18; Gaps 2;

QY 37 EIANFVRPSDVWIVTPKSGTSLLOEVVYLVNSQADPEIGLMNIDQLPVLEYQPGSL 96
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 30 QVESFEARPDILISTYPKSGTTWISILDLYNNGDAEKCCKRDAIYKRVFPMELIIPGI 89
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 97 ----DIKELTSPRLKSHLPYRFLPSDLHNGDSKVIYMARNPDKLVVSYQFHRSLRTM 152
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 90 TNGVEMLNMPSPRIVKTHLPVQLPSPSFWKNDCKIIVARNAKDVVSYFYQMAKIH 149
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 153 SYRGTFQFCRCRRFMNDKLGYSWFHVFQFWEHRMDSNVLFKYEDMHRDLVTWVQLAR 212
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 150 PEPGTWEEFLEKFMAGVSPGPDYDHVKSWKEKRYRILYLFYEDMKENPKCEIQKILK 209
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 213 FLGVSCDKAQLA-LTEHCH-----QLVDOCCNAELPVGRGRVGLWKDITF 258
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 210 FLEKDIPEILNKILYHSSFSVMKNPANSYTTMMKEEMDSVSPFMKRGISGDMKNQFT 269
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 259 VSMNEKFDLVYKQMGKCDLTF 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 270 VAQYEKFEEDYVKMGKESTLKF 291
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 86
Q66KW4 XENLA
ID Q66KW4_XENLA PRELIMINARY; PRT; 305 AA.
AC Q66KW4;
DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE MGC85375 protein.
GN Name=MGC83375;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney P.H.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

```

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Whole;  
 RA Klein S., Gerhard D.S.;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; BC078536; AAH78536.1; -; mRNA.  
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.  
 DR InterPro; IPR000863; Sulfotransferase.  
 DR Pfam; PF00685; Sulfotransferase 1; 1.  
 DR ProDom; PD001218; Sulfotransferase; 1.  
 SQ SEQUENCE 305 AA; 35569 MW; PCS175F6S30564DF CRC64;  
 Query Match 26.9%; Score 411; DB 2; Length 305;  
 Best Local Similarity 33.7%; Pred. No. 7.3e-26;  
 Matches 101; Conservative 48; Mismatches 123; Indels 28; Gaps 6;  
 QY 1 MAESAETPTPGFESKVFEGHVRP-PFCRGKMEIANPVPVDPVWVITYPKSGTS 59  
 Db 9 LAEEMENFQVTMGHIE-----GVPLPGPTC-DEWDTIYFQAREDDILIATYPKSGTT 60  
 QY 60 LLOEVVYLVSQADPDEIGLNIQDLVLEYPQP-----GLDIITKELTSPRLIKSHLPY 114  
 Db 61 WMOEIVDLIQEGDVQKSWRAPCFIKVFIEMIPKSPMSGLAKTWKSPRLIKTHLP 120  
 QY 115 RFLPSDLHNGSKVIYMARNPDKLVVSYQFHRSLRTWSYRGTFQFCRRFMNDKLGYS 174  
 Db 121 NLLPFSFEKNAKVYVARNKDVVYFHKMTFLDLPCTWDFNFEFLSGDVPWGS 180  
 QY 175 WFEHQVFEHRMDSNVFLKYEDMHRDLVTWVEQLARFLGVSCKAQLEALTEHCH--- 231  
 Db 181 WFDHVLGWMKMDKHQILFIYEDIMEDPMREIRKVMKFLGKDLSDALEKNVYHSSFQA 240  
 QY 232 QLVDQCCNAEALP-----VGRGRVGLWKDIFTVMNEKFDLVYKQMGKCDLTF 280  
 Db 241 MKENPMTNSTVPSNIMDDTISPFMRKGLVGDWKTHFSVTQNFIDKEYKKMGESGLNF 300  
 RESULT 87  
 ID Q5HZUO\_XENTR PRELIMINARY; PRT; 297 AA.  
 AC Q5HZUO;  
 DT 15-FEB-2005, integrated into UniProtKB/TrEMBL.  
 DT 15-FEB-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 10.  
 DE Hypothetical LOC496998.  
 GN Name=LOC496998;  
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; Xenopus; Silurana.  
 OX NCBI\_TaxID=8364;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Whole body;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Whole body;  
 RA Klein S., Gerhard D.S.;  
 RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; BC088888; AAH88888.1; -; mRNA.  
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.  
 DR InterPro; IPR000863; Sulfotransferase.  
 DR Pfam; PF00685; Sulfotransferase 1; 1.  
 DR ProDom; PD001218; Sulfotransferase; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 297 AA; 34863 MW; B8248162DC31D7D2 CRC64;  
 Query Match 26.8%; Score 410.5; DB 2; Length 297;  
 Best Local Similarity 32.4%; Pred. No. 7.7e-26;  
 Matches 85; Conservative 61; Mismatches 99; Indels 17; Gaps 4;  
 QY 34 KMEIANPVPVDPVWVITYPKSGTSLLOEVVYLVSQADPDEIGLNIQDLVLEYP 92  
 Db 30 KIDSTQDFKVDVTLVFTYPTKTQITWQILSLIFNEGHRNGTEAIANV-FRVPWIEYT 88  
 QY 93 QPGDLIKELTSPRLIKSHLPYRFLPSDLHNGSKVIYMARNPDKLVVSYQFHRSLRTM 152  
 Db 89 HSKVD-YDSRSPSPFLFSSHLPHYLVPKDLRNKKKIIYVGRNPKDAVSYHYFNVIRL 147  
 QY 153 SYRGTFQFCRRFMNDKLGYSWFEHQVFEHRMDSNVFLKYEDMHRDLVTWVEQLAR 212  
 Db 148 KQVNDWESFLDYLTLGVLGGSWFDHVKGWYTHQEDFNILFVTEEMKDLRSALVKICK 207  
 QY 213 FLGVSCKAQLEALTE-----HCHQLVDQCCNAEALPVGGRVGLWKDIFT 258  
 Db 208 FVEKELNEQEVDTIVERKATFKMKHDPPLANVTNTVDHLMKNGTFLRGTGVDKMKLMT 267  
 QY 259 VSMNEKFDLVYKQMGKCDLTF 280  
 Db 268 VAQNEKFDKIYSEKMGVPINF 289  
 RESULT 88  
 ID STIE1\_CAVPO STANDARD; PRT; 296 AA.  
 AC P49887;  
 DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.  
 DT 01-OCT-1996, sequence version 1.  
 DT 07-FEB-2006, entry version 43.  
 DE Estrogen sulfotransferase (EC 2.8.2.4) (Sulfotransferase, estrogen-  
 preferring) (STIE3).  
 DE Name=SULT1E1; Synonyms=STE;  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 OC Hystricognathi; Caviidae; Cavia.  
 OX NCBI\_TaxID=10141;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [MRNA], AND PARTIAL PROTEIN SEQUENCE.  
 RC STRAIN=NH 2; TISSUE=Adrenal cortex;  
 RX MEDLINE=93024479; PubMed=1406700; DOI=10.1210/me.6.8.1216;  
 RA Oeda T., Lee Y.C., Driscoll W.J., Chen H.-C., Strott C.A.;  
 RT "Molecular cloning and expression of a full-length complementary DNA

```

RT encoding the guinea pig adrenocortical estrogen sulfotransferase.";
RL Mol. Endocrinol. 6:1216-1226(1992).
RN [2]
RP PAPS-BINDING SITE, AND MUTAGENESIS.
RA MEDLINE=95071451; PubMed=7980593;
RX Komatsu K., Driscoll W.J., Koh Y., Strott C.A.;
RT "A p-loop related motif (GxGxxK) highly conserved in
RT sulfotransferases is required for binding the activated sulfate
RT donor.";
RL Biochem. Biophys. Res. Commun. 204:1178-1185(1994).
CC -!- FUNCTION: Sulfation of estradiol and estrone. May control the
CC level of the estrogen receptor by sulfonylating free estradiol.
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =
CC adenosine 3',5'-biphosphate + estrone 3-sulfate.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- TISSUE SPECIFICITY: Adrenal gland and much less in liver. Uterine
CC STE is detectable only during pregnancy.
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U09552; AAA18495.1; -; mRNA.
DR PIR; A44011; A44011.
DR SMR; P49888; IG3M.
DR HSP; P49887; 5-294.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; P0001218; Sulfotransferase; 1.
KW Direct protein sequencing; Lipid-binding; Steroid-binding;
KW Transferase.
FT CHAIN 1 296 Estrogen sulfotransferase.
FT FTIG-PRO 0000085152.
FT NP_BIND 49 54 PAPS (By similarity).
FT NP_BIND 131 139 PAPS (By similarity).
FT NP_BIND 194 230 PAPS (By similarity).
FT NP_BIND 238 260 PAPS (By similarity).
FT ACT_SITE 109 109 Proton acceptor (By similarity).
FT MUTAGEN 260 260 G->A: Loss of activity.
FT MUTAGEN 263 263 G->A: Loss of activity.
FT MUTAGEN 266 266 K->A: Loss of activity.
SQ SEQUENCE 296 AA; 35245 MW; 019D92B95E29901D CRC64;

Query Match 26.7%; Score 408.5; DB 1; Length 296;
Best Local Similarity 34.2%; Pred. No. 1.1e-25;
Matches 94; Conservative 50; Mismatches 112; Indels 19; Gaps 4;

QY 18 KYF-EFHGVLRLPPFCRGKMEETANFVRPSDVWIVTYPKSGTSLLEQVYLVYSGQADPDE 76
DB 11 EYFDFRGILLYKQIKYKNDVNEAFQARPDDLVIAAPKSGTITWSEVVCMIYAGDVKK 70

QY 77 IGLMNIDEOLPVLVEYPPQ-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMA 132
DB 71 CQDAIFPNRVPLECNDKMGVQLEEMSPRIKTHLPRLLPASFWEKRCWKICIC 130

QY 133 RNPDLVSVYQFHRSRLTMSYRGTFQBFRRFMNDKLGYSQWGFHVFWEHRMDSNVL 192
DB 131 RNAKDVAVSYYYFFLVMANHPDGSFPFVEKFMQGVPGYQWYDHVKSWEKSTDPRL 190

QY 193 FLKYEDMRDLVTVMVEQLARFLGVSCKAQLEALTEHCH-----QLVDOCN 239
DB 191 FIFEDMKEDIRKEVLKHLHFGRKPSBELVDKIHKTSFQEMKNPNSTNYTMLPEEIMN 250

QY 240 AEALP-VGRGRVGLWKDIFTVSMNEKFDLVYQKM 273
DB 251 QKVSFPMKGISGDWKNHFTVALNESFDKHYQQM 285

RESULT 89
O70262_MOUSE PRELIMINARY; PRT; 304 AA.
ID O70262_MOUSE

```

```

AC O70262;
DT 01-AUG-1998, integrated into UniProtKB/TrEMBL.
DT 01-AUG-1998, sequence version 1.
DT 07-FEB-2006, entry version 26.
DE Phenol sulfotransferase.
GN Name=Sult1c1; Synonyms=Sult1a2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ddv; TISSUE=Olfactory;
RX MEDLINE=96208491; PubMed=8641270;
RA Miyawaki A., Homma H., Tamura H., Matsui M., Mikoshiba K.;
RT "Zonal distribution of sulfotransferase for phenol in olfactory
RT sustentacular cells.";
RL EMBL J. 15:2050-2055(1996).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ddv; TISSUE=Olfactory;
RX MEDLINE=98228237; PubMed=9560327;
RA Tamura H.-O., Harada Y., Miyawaki A., Mikoshiba K., Matsui M.;
RT "Molecular cloning and expression of a cDNA encoding an olfactory-
RT specific mouse phenol sulphotransferase.";
RL Biochem. J. 331:953-958(1998).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ddv; TISSUE=Olfactory;
RA Tamura H.-O., Harada Y., Miyawaki A., Mikoshiba K., Matsui M.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF033653; AAC17740.1; -; mRNA.
DR HSP; P50224; 1CJM.
DR EMBL; ENSMUSG0000003943; Mus musculus.
DR MGI; MGI:102928; Sult1c1.
DR GO; GO:0005737; Cytoplasm; IDA.
DR GO; GO:0004062; P:aryl sulfotransferase activity; IDA.
DR GO; GO:0006790; P:sulfur metabolism; IDA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; P0001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 304 AA; 35797 MW; 30857261A1462BE4 CRC64;

Query Match 26.7%; Score 408.5; DB 2; Length 304;
Best Local Similarity 31.5%; Pred. No. 1.2e-25;
Matches 88; Conservative 57; Mismatches 115; Indels 19; Gaps 3;

QY 21 EFHGVLRLPPFCRGKMEETANFVRPSDVWIVTYPKSGTSLLEQVYLVYSGQADPDEIGLM 80
DB 22 EVNGILSMKMSNWNKLNWFQAKPDDLIIATYAKAGTTWTQEIIVDMIQNDGVDQKQRA 81

QY 81 NIDEOLPVLVEYPPQ-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
DB 82 NTDYRHPFIETWLPPLNSGLDLANKMPSRPTKTHLPVQMLPPSPFWKENSQIIYVARNA 141

QY 136 KDLVVSVYQFHRSRLTMSYRGTFQBFRRFMNDKLGYSQWGFHVFWEHRMDSNVLFK 195
DB 142 KDLVSVYFYFRMKNKLPDPDTGLGEIYTFKAGKVLGWSWDVHVGWWDVKKRILYLF 201

QY 196 YEDMRDLVTVMVEQLARFLGVSCKAQLEALTEHCHQLV---DQCCNAELP----- 244
DB 202 YEDMKDPKEIKKIVKLEKDISSEVLNKKIHHSTFDMKQNPANVTTLTSSIMDHSI 261

QY 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYQKMCKDLTF 280
DB 262 SPFMKGMKGMKKNYFTVAQSEDFEDYRKKMAGSTITF 300

```



```
RESULT 90
Q4RG19 TETNG
ID Q4RG19 TETNG PRELIMINARY; PRT; 309 AA.
AC Q4RG19;
DT 19-JUL-2005, integrated into UniProtKB/T-EMBL.
DT 19-JUL-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Chromosome 2 SCAP15106, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSTENG0030505001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Maucell E., Bounneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.,
RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype."
RL Nature 431:946-957 (2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonCommercial License
CC -----
DR EMBL; CA01015106; CAG12663.1; -; Genomic DNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
FT NON TER 309 309
SQ SEQUENCE 309 AA; 36213 MW; E60810D9271AFFF14 CRC64;

Query Match 26.5%; Score 405.5; DB 2; Length 309;
Best Local Similarity 33.7%; Pred. No. 2.1e-25;
Matches 93; Conservative 49; Mismatches 115; Indels 19; Gaps 4;

QY 24 GVRLPPFCRGMEEIANFVRPSDVIMVTPKSGTSLQEVYLVYSGQADPDEIGLMNID 83
DB 30 GVPLMSHTAQNFDISWAFRPPDSLLIATYPAGTTWTQEIYVDLLHNGDADACKRPTP 89

QY 84 EQLPVLEY--POP---GLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVIYMARPKDL 138
DB 90 VRSPELEYAPPPIPSGLDLKXNDPPRPFKTHLPQLVPPAPFAPWENKCKTIYVARNAKDN 149

QY 139 VVSYVQFHRSLRTMSYRGTFQFCFRFNMNKLGYCSWPHQVFEHRMDSNVLFKYED 198
DB 150 LVSYVFFDCMNMNTPQEPGTMEEYIHKFMRGELSGWSGDVHVYKWKDKNKNILFYED 209

QY 199 MHRDLVTWVQLARFLVGSCKD---AQLEALTEHCQLVDQCCNAEALP----- 244
DB 210 MKENPRPREVERIMRVLDVSVSDEVISKVILTSFPEKMKENPMNATYCPAPVDFHSKSPF 269
```

```
QY 245 VGRGRVGLWKDIFTVMNKEKFDLVTKQKMGKCDLTF 280
DB 270 MRKGQVGRNRYFTPEQKMFEEYKQKQKVDIIF 305

RESULT 91
Q9VHH0 DROME
ID Q9VHH0 DROME PRELIMINARY; PRT; 316 AA.
AC Q9VHH0;
DT 01-MAY-2000, integrated into UniProtKB/T-EMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 27.
DE CG16733-PA.
GN ORFNames=CG16733, Dmel_CG16733;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalaali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195 (2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleeb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).
RN [3]
```

RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirskas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celisner S.E.;  
RT "The transposable elements of the Drosophila melanogaster euchromatin:  
RT a genomic perspective";  
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Muncall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bettencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yanada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
RT systematic review";  
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
RN [5]  
RP NUCLEOTIDE SEQUENCE.  
RG Berkeley Drosophila Genome Project;  
RA Celisner S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,  
RA Hoskins R., Stapleton M., Facleb J., Park S., Svirskas R., Smith E.,  
RA Yu C., Rubin G.;  
RT "Drosophila melanogaster release 4 sequence";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP NUCLEOTIDE SEQUENCE.  
RG FlyBase;  
RL Submitted (JAN-2006) to the EMBL/GenBank/DBJ databases.  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; AE003682; AAF54344.1; -; Genomic\_DNA.  
DR HSP; P49888; IHV3.  
DR FlyBase; FBgn0037665; CG16733.  
DR GO; GO:0008146; F:sulfotransferase activity; IEA.  
DR InterPro; IPR000863; Sulfotransferase.  
DR Pfam; PF00685; Sulfotransfer 1; 1.  
SQ SEQUENCE 316 AA; 37332 MW; EBCD9D561E4E1C77 CRC64;  
  
Query Match 26.5%; Score 405; DB 2; Length 316;  
Best Local Similarity 34.4%; Pred. No. 2.4e-25;  
Matches 94; Conservative 52; Mismatches 99; Indels 28; Gaps 8;  
  
QY 36 EBIANFPVPSVWIVTPYKSTSLQEVVYLSQ-----GADPD-----EI-GLMN 81  
DB 45 ESIRSLPVYQDDVMVSPRTGSTWAQENVMVLGHQLDYVAEQDLRLSLPIELSLF 104  
QY 82 IDEQLPVLEYPOGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVITYMARNPKDLVVS 141  
DB 105 IDHHTVTAQFGNTVDLVNLRPRFARSHLPWLLPEQFETVKPRIVYIATARNPKDLVVS 164  
QY 142 YYQFHRSLRTMSYRGTFQFCRRFRMNDKLGYSWFHVEHQEFWEHRMDSNVLFKYEDMHR 201  
DB 165 VYHYFKLLHGMN--GDPEQFVDFLEGHTPMGSYWRHVLPPFKRSQDDNVLFKYEDVMK 222  
QY 202 DLVTVWEQLARFLGVS--CDKAQLALTEHC---HQLVDCCNAE-ALP-----VGRGR 249  
DB 223 DLPSVVRRCARFLGVSQSLDVSSTLQKLDHLTFDKMRANKAIVLEKLLPSESSKFIIRNGK 282  
QY 250 VGLWKDIFTVMNKEKFDLVYKMGKCDLTFDF 282  
DB 283 IGDRNHNHGMNSERFDEWTERHMRGSGLNFDY 315  
  
RESULT 92  
Q6DJL6\_XENLA  
ID Q6DJL6\_XENLA PRELIMINARY; PRT; 287 AA.

AC O6DJL6;  
DT 16-AUG-2004, integrated into UniProtKB/TrEMBL.  
DT 16-AUG-2004, sequence version 1.  
DT 07-FEB-2006, entry version 14.  
DE MGC82053 protein.  
GN Name=MGC82053;  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus; Xenopus.  
OX NCBI\_TaxID=83355;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Kidney;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Hulyk S.W.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Kidney;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.;  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative";  
RL Dev. Dyn. 225:384-391(2002).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Kidney;  
RA Klein S., Gerhard D.S.;  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; BC075160; AAH75160.1; -; mRNA.  
DR GO; GO:0008146; F:sulfotransferase activity; IEA.  
DR InterPro; IPR002202; HMG-CoA\_red.  
DR InterPro; IPR000863; Sulfotransferase.  
DR Pfam; PF00685; Sulfotransfer 1; 1.  
DR ProDom; PD001218; Sulfotransferase; 1.  
DR PROSITE; PS01192; HMG COA REDUCTASE 3; UNKNOWN 1.  
SQ SEQUENCE 287 AA; 33812 MW; EB655BD8D29B0AB7 CRC64;  
  
Query Match 26.4%; Score 404.5; DB 2; Length 287;  
Best Local Similarity 30.9%; Pred. No. 2.4e-25;  
Matches 85; Conservative 57; Mismatches 116; Indels 17; Gaps 3;  
  
QY 23 HGVRLPPFCRGKMEIANFPVPSVWIVTPYKSTSLQEVVYLSQADPDEGLMNI 82  
DB 9 NGVPLSEAIASNDWEIWFQAKPGDILINTPKSGTTWQEIILIMDVDEKCKRAPI 68  
QY 83 DEQLP---VLEYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVITYMARNPKDLV 139  
DB 69 YERIPFVDILHLMKPGVEVNAMPSPRLIKSHLAFQLVPPSPFWKECKVIVYARNAKDTA 128

```

Qy      140 VSYQYHRSRLRTWSVRGTFQECRFRPMNDKLGYSWFHFEHQESFWEHHRMDSNVFLKYEDM 199
Db      129 TSFFYFDMAHLHPTPEGWDDYLERFMFGDVGWGSYDHWKGFWEQKQDNILYLFEDL 188
Qy      200 HRDLVTMVQLARLVGSCDQAQLEA---LTHECHQLVQCCNAEALP-----V 245
Db      189 QKNPLVEIRKVMKFLDKOLPEELRLKIVHLTSPKKMKENPMANYSTFSPDMLAQTDHKFM 248
Qy      246 GGRGVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db      249 RKGKVGDMKQFTVYQNEVFADYQRMQSGSTLKF 283

```

RESULT 93	
Q5W7C2_XENLA	
ID	Q5W7C2_XENLA PRELIMINARY; PRT; 287 AA.
AC	Q5W7C2;
DT	01-FEB-2005, integrated into UniProtKB/TrEMBL.
DT	01-FEB-2005, sequence version 1.
DT	07-FEB-2006, entry version 10.
DE	LOC436245 protein.
GN	Names=LOC436246;
OS	Xenopus laevis (African clawed frog).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC	Xenopodinae; Xenopus; Xenopus.
OX	NCBI_TaxID=8355;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RC	TISSUE=Testis;
RX	MEDLINE=223411132; PubMed=12454947; DOI=10.1002/dvdy.10174;
RA	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA	Richardson P.,
RT	"Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT	Initiative.";
RL	Dev. Dyn. 225:384-391(2002).
RN	[2]
RP	NUCLEOTIDE SEQUENCE.
RC	TISSUE=Testis;
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA	Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Usdin T.B., Tshiyuki S., Carninci P., Prange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA	Bosak S.A., McWain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Vallalon D.K., Morzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fahy J.J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA	Butterfield F.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences.";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN	[3]
RP	NUCLEOTIDE SEQUENCE.
RC	TISSUE=Testis;
RA	Klein S., Gerhard D.S.;
RL	Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC	
CC	Copyrighted by the UniProt Consortium, see <a href="http://www.uniprot.org/terms">http://www.uniprot.org/terms</a>
CC	Distributed under the Creative Commons Attribution-NoDerivs License
CC	
CC	EMBL; BC088717; AAH88717.1; -; mRNA.
DR	GO; GO:0008146; F:sulfotransferase activity; IEA.
DR	InterPro; IPR000863; Sulfotransferase.

```

DR Pfam: PF00695; Sulfotransfer 1; 1.
DR ProDom: PD001219; Sulfotransferase; 1.
SQ SEQUENCE 287 AA; 34129 MW; 138C774D7CE4A591 CRC64;

Query Match          26.3%; Score 402; DB 2; Length 287;
Best Local Similarity 30.9%; Pred. No. 3.8e-25;
Matches 88; Conservative 104; Indels 26; Gaps 6

Qy 19 YFPHGVRLLP--FCRCGMBEIAN-FVPRPSDVMIVTPKSGTSLLOEYVVLVSQGDPD 75
   |||: |||: : : : : |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 5 YFLKGVKSPGSGYQELLQVYVNEFQVLDVVIVTPKSGTHWLIELSILNRDADPT 64
   |||: |||: : : : : |||: |||: |||: |||: |||: |||: |||: |||: |||:

Qy 76 EIGLMTDEQLPVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
   |||: |||: : : : : |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 65 WCNSVLTLWRSPWITKEGQLQ-LKOLSSPRLVTLSHLPFIFFPKAFFTSKAKIIYVMRNP 123
   |||: |||: : : : : |||: |||: |||: |||: |||: |||: |||: |||: |||:

Qy 136 KDLVVSYYQHRSLRTWSYRGTFQFCRRFMNDKLGYSWFHGVQEFWEHRMDSNVFLK 195
   |||: |||: : : : : |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 124 KDIFVSLFYFAKICHYKDPSPFQYLEDFLQGNILYNSWFHVRGWMQKDNSNFFIIT 183
   |||: |||: : : : : |||: |||: |||: |||: |||: |||: |||: |||: |||:

Qy 196 YEDMHRDLRTVMEQLARFLVGCSDKAQLEALTEH-----CHQLVDOCC 238
   |||: |||: : : : : |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 184 YEELHQDLRGCIILRICFLKELGDSKIDTIKVHSSFEAMKENKMSNYSLGTTEEYID--- 240
   |||: |||: : : : : |||: |||: |||: |||: |||: |||: |||: |||: |||:

Qy 239 NASALPVGRGRVGLWKLQIFTVSMNNEKFDLVYKQMGKCDLTFDFY 283
   |||: |||: : : : : |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 241 HTKGTGFMRKGTAGDKWNKHFVVAOSEHFDVRYQEKVK--DLNNKFF 283
   |||: |||: : : : : |||: |||: |||: |||: |||: |||: |||: |||: |||:

```

RESULT	94
C491K6_ERAARE	PRELIMINARY; PRT; 293 AA.
ID	Q491K6_BRARE
AC	Q491K6;
DT	13-SEP-2005, integrated into UniProtKB/TrEMBL.
DC	13-SEP-2005, sequence version 1.
DD	07-FEB-2006, entry version 2.
DE	SULT1 isoform 5.
DN	Name=sult1st5;
OS	Brachydanio rerio (Zebrafish) (Danio rerio).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC	Cyprinidae; Danio.
OX	NCBI_TaxID=7955;
RN	[1]
RP	NUCLEOTIDE SEQUENCE
RF	PubMed:16045754; DOI=10.1111/j.1742-4558.2005.04791.x;
RA	Yasuda S., Kumar A.P., Liu M.Y., Sakakibara Y., Suiko M., Chen L.,
RA	Liu M.C.;
RT	"Identification of a novel thyroid hormone-sulfating cytosolic
RT	sulfotransferase, SULT1 ST5, from zebrafish.";
RL	PBES J. 272:3828-3837(2005).
RN	[2]
RP	NUCLEOTIDE SEQUENCE.
RF	Liu M.-C.;
RL	Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC	Distributed under the Creative Commons Attribution-NoDerivs License
CC	-----
CC	EMBL; AY879099; AAY47051.1; -; mRNA.
DR	ZFIN; ZDB-GENE-050809-2; sult1st5.
DR	GO; GO:0008146; F:sulfotransferase activity; IDA.
DR	GO; GO:0006805; P:xenobiotic metabolism; IDA.
DR	InterPro; IPR000863; Sulfotransferase.
DR	Pfam; PF00685; Sulfotransferase 1; 1.
DR	ProDom; PD001218; Sulfotransferase; 1.
SQ	SEQUENCE 293 AA; 34454 MW; 108BB08C28C4A3BC CRC64;
Query Match	26.2%; Score 401.5; DB 2; Length 293;
Best Local Similarity	32.6%; Pred. No. 4.3e-25;
Matches	93; Conservative 44; Mismatches 121; Indels 27; Gaps





```
Db 158 LPDPTGTDYFMAFLSGTLPMGWSFHDVNGWAKDKHRVLVYFYEDMKKOLRLRIQRVE 217
Qy 212 RFLGVSCDKAQLEALTEH-CHQLVQ--CCNAELP-----VGRGRVLGWKIDIF 257
Db 218 SFPLDKDLPEEVLKICQHTTFQAMKENPMANYTTPTTMDQSPFPMRKGVGDWKNHF 277
Qy 258 TVSMNEKFDLVYKQKMGKCDLTF 280
Db 278 LVAQNELFDWEYKRRMDGTGLDF 300

RESULT 99
Q7ZYHO_XENLA
ID Q7ZYHO_XENLA PRELIMINARY; PRT; 304 AA.
AC Q7ZYHO;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 17.
DE Sultic1-prov protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus;
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., McEwan P.J., McKernan K.J., Abramson R.D., Mullahy S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; BC043790; AAH43790.1; -; mRNA.
DR HSP; P50224; 1CJM.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
SQ SEQUENCE 304 AA; 35108 MW; 0D4C0C2B017C5692 CRC64;
```

```
Query Match 25.7%; Score 393.5; DB 2; Length 304;
Best Local Similarity 33.1%; Pred. No. 2.1e-24;
Matches 95; Conservative 49; Mismatches 120; Indels 23; Gaps 5;

Qy 15 PESKYFEPFHGVRLPPFCRGMKEEIANPVRSDVIMVITPKSGTSLLOEVVLVSQGAD- 73
Db 16 FQVTWGHIEGVPLPVTTCDMDTINFAQKDDILIIATYPKAGTTWQEIVDLILQEGDV 75
Qy 74 -----PDEIGLMNIDQLPVLEYQPQGLDIIKELTSPRLIKSHLPFLFSLDHLNGDSK 127
Db 76 QKSMRAPCYIKVPFID-LVPPKPMP-DGVALAQMTNSPRLKTHLPINLLPPEFWKXNTK 133
Qy 128 VIYNARNPKDLVVSYYQFHSRLRTWSYRGTFQECRRPMDKLGYSWFHVQSFWEHRM 187
Db 134 VVYVARNAKDSMVSYYPHKNKFLPSDGLTLDNFSEFLSGDVPWGSFDFNLVGMWKALD 193
Qy 188 DSNVFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLEALTEHCH---QLVQCCNAELP 244
Db 194 KQIQLFIYEDMIQDPHREIKKVTFLGKDLSDLEVIKLYHTSFQAKENPMTNNSTVP 253
Qy 245 -----VGRGRVLGWKIDFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 254 KTIMDQTISPRIKRGTVGDWKTFFSVAQNIIFDEYKKNKEGSLNF 300

RESULT 100
Q6GL93_XENTR
ID Q6GL93_XENTR PRELIMINARY; PRT; 299 AA.
AC Q6GL93;
DT 19-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE MGC69544 protein.
GN Name=MGC69544;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC
```



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2006, 11:17:09 ; Search time 10 Seconds  
(without alignments)  
60.666 Million cell updates/sec

Title: US-10-768-158-2

Perfect score: 1530

Sequence: 1 MAESEATPTGFESEKYP.....FDLVYKQKMGKDLTFDPYL 284

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 21570 seqs, 2136119 residues

Total number of hits satisfying chosen parameters: 21570

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications AA.New.\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	77.5	5.1	411	6	US-10-505-928-445
2	75	4.9	337	6	US-10-196-749-268
3	75	4.9	337	7	US-11-101-316-74
4	73.5	4.8	273	6	US-10-511-814-1
5	72.5	4.7	365	7	US-11-251-465-25
6	72.5	4.7	367	7	US-11-312-958-14
7	72	4.7	813	6	US-10-196-749-466
8	71	4.6	259	7	US-11-249-937-2407
9	69	4.5	1534	6	US-10-505-928-40
10	68.5	4.5	634	7	US-11-311-778-18
11	68.5	4.5	820	6	US-10-982-908-26
12	67.5	4.4	424	7	US-11-242-505A-33
13	67.5	4.4	1056	6	US-10-505-928-225
14	65	4.2	587	6	US-10-511-937-2511
15	64.5	4.2	2202	6	US-10-488-015-12
16	64	4.2	226	7	US-11-183-218-42
17	63.5	4.2	209	6	US-10-196-749-270
18	63.5	4.2	226	6	US-10-984-694-3
19	62.5	4.1	226	6	US-10-984-694-1
20	62.5	4.1	933	6	US-10-511-455-50
21	62.5	4.1	956	6	US-10-511-455-41
22	62	4.1	309	7	US-11-283-522-34
23	62	4.1	492	6	US-10-505-928-642
24	62	4.1	1085	6	US-10-505-928-343
25	62	4.1	1722	6	US-10-505-928-780

Sequence 5, Appli  
Sequence 232, App  
Sequence 2512, Ap  
Sequence 40, Appl  
Sequence 2584, Ap  
Sequence 772, App  
Sequence 37, Appl  
Sequence 30, Appl  
Sequence 35, Appl  
Sequence 140, App  
Sequence 404, App  
Sequence 124, App  
Sequence 22, Appl  
Sequence 351, App  
Sequence 19, Appl  
Sequence 175, App  
Sequence 401, App  
Sequence 3, Appli  
Sequence 644, App  
Sequence 304, App  
Sequence 449, App  
Sequence 496, App  
Sequence 2, Appli  
Sequence 662, App  
Sequence 2, Appli  
Sequence 646, App  
Sequence 416, App  
Sequence 21, Appl  
Sequence 288, App  
Sequence 491, App  
Sequence 35, Appl  
Sequence 100, App  
Sequence 167, App  
Sequence 2991, Ap  
Sequence 4, Appli  
Sequence 7, Appli  
Sequence 54, Appl  
Sequence 4, Appli  
Sequence 17, Appl  
Sequence 336, App  
Sequence 82, Appli  
Sequence 2, Appli  
Sequence 11, Appl  
Sequence 56, Appl  
Sequence 3, Appli  
Sequence 8, Appli  
Sequence 40, Appl  
Sequence 145, App  
Sequence 84, Appl  
Sequence 803, App  
Sequence 393, App  
Sequence 1, Appli  
Sequence 21, Appl  
Sequence 729, App  
Sequence 14, Appl  
Sequence 1903, Ap  
Sequence 2546, Ap  
Sequence 35, Appl  
Sequence 2511, Ap  
Sequence 13, Appl  
Sequence 152, App  
Sequence 2413, Ap  
Sequence 1390, App  
Sequence 366, App  
Sequence 741, App  
Sequence 12, Appl  
Sequence 1819, Ap  
Sequence 359, App  
Sequence 24, Appl  
Sequence 10, Appl  
Sequence 12, Appl

99 54.5 3.6 1178 7 US-11-311-778-14 Sequence 14, Appl  
100 54.5 3.6 1575 6 US-10-505-928-257 Sequence 257, App

## ALIGNMENTS

## RESULT 1

US-10-505-928-445  
; Sequence 445, Application US/10505928  
; Publication No. US20060088532A1  
; GENERAL INFORMATION:  
; APPLICANT: Ludwig Institute for Cancer Research et al.  
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES  
; FILE REFERENCE: 28967/39178  
; CURRENT APPLICATION NUMBER: US/10/505,928  
; CURRENT FILING DATE: 2004-08-27  
; PRIOR APPLICATION NUMBER: US 60/363,019  
; PRIOR FILING DATE: 2002-03-07  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: PatentIn 3.2  
; SEQ ID NO 445  
; LENGTH: 411  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-505-928-445

Query Match 5.1%; Score 77.5; DB 6; Length 411;  
Best Local Similarity 17.3%; Pred. No. 0.5;  
Matches 51; Conservative 56; Mismatches 84; Indels 103; Gaps 13;  
QY 19 YPEFHGVRLPP-----FCGKMEETANFPV-----RPSDVWIVTPKSGTSLLEQVYV 66  
Db 133 YFLENYIKPPVNNHTTDRIFRGGASRVLCSPVCDPPGPADL----- 174  
QY 67 LVSQADPDEIGLMDID-----EQLPVLEYQPQGLDIKELT-SPRLIKSHLPYRF 116  
Db 175 VLEEGDCVRKCGLLNLTAEEACRERSHVAIKTVRVPENDURALVEDPRL----- 225  
QY 117 LPSDLHGSDKVIYMARNPDKDVLVSYQFHRSLRTMSVGTGFQFCFRFMNDKLGYG--- 173  
Db 226 -----NLKVIQLVRDPRGILAS-----RSETFRDTYRLM-----RLWYGTGR 262  
QY 174 -----SWFEHVQGFWEHRMDSNV-----LFLKYEDMHRDLVTMVEQLARFLGV 216  
Db 263 KYPNLDVTQLTTVCBDFNSVSTGLMRPWLKGYMLVRYEDLARNPMKKTBEIYGFGLGI 322  
QY 217 SCDKAQLREALTECHQLVDQCCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVY 269  
Db 323 PLD-----SHVARWIQN--NTRGDDPTLGHKKHGTVRNSAATAEKWRFRLSY 366

## RESULT 2

US-10-196-749-268  
; Sequence 268, Application US/10196749  
; Publication No. US20060094864A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430RIC340  
; CURRENT APPLICATION NUMBER: US/10/196,749  
; CURRENT FILING DATE: 2002-07-16

; PRIOR APPLICATION NUMBER: 10/052586  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059266  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063120  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063121  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063486  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/063540  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063541  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063544  
; PRIOR FILING DATE: 1997-10-28  
; Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 268  
; LENGTH: 337  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-196-749-268  
Query Match 4.9%; Score 75; DB 6; Length 337;  
Best Local Similarity 18.8%; Pred. No. 0.7;  
Matches 33; Conservative 16; Mismatches 41; Indels 86; Gaps 7;  
QY 47 DWIVTVTPKSGTSLLEQV-VVLVSQADPDPBIGLMDIQLPVLE---YPOPGLDIILKEL 102  
Db 31 DHPASYPECCNNAQSPIDIQTDSVTFDPD-----LPALQPHGYDQPGTE----- 75  
QY 103 TSPRLIKSHLPYRFLPDLHNGDSKV-----TYMARNPDKDVLVSYQFHRSLRTMSYR 155  
Db 76 -----PLDHNHGHVTVQLSLPSTVLGLPKRYVAAQLHLH----- 111  
QY 156 GTFQBFCCRFRMNDKLGYSWFEHVQEFW-----EHRMDSNVFLFLKYEDMHRD 202  
Db 112 -----WGKGSGSGSEHQINSEATFAELHIVHYD 140  
RESULT 3  
US-11-101-316-74  
; Sequence 74, Application US/1101316  
; Publication No. US20060099657A1  
; GENERAL INFORMATION:  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID  
; FILE REFERENCE: P3230RIC17C1  
; CURRENT APPLICATION NUMBER: US/11/101,316  
; CURRENT FILING DATE: 2005-04-06  
; PRIOR APPLICATION NUMBER: 10/063526  
; PRIOR FILING DATE: 2002-05-03  
; PRIOR APPLICATION NUMBER: 10/006867  
; PRIOR FILING DATE: 2001-12-06  
; PRIOR APPLICATION NUMBER: PCT/US00/23328  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 09/380137  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: PCT/US99/12252  
; PRIOR FILING DATE: 1999-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; NUMBER OF SEQ ID NOS: 170

```
; SEQ ID NO 74
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-101-316-74

Query Match          4.9%; Score 75; DB 7; Length 337;
Best Local Similarity 18.8%; Pred. No. 0.7;
Matches 33; Conservative 16; Mismatches 41; Indels 86; Gaps 7;

Qy 47 DWIVTPKSGTSLQEV-VVLVSGADPPDEIGLNMIDEQLPVLE---YPOPLGDIKEL 102
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 31 DHWPASYECGNAQSPIDIQDTSVTFDPD-----LPALOPHYGDQPGTE----- 75

Qy 103 TSPRLIKSHLPYRFLPSDLHNGDSKV-----IYMARNPDLVSVYYQFHRSLRTMSYR 155
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 76 -----PLDLHNNGHVTQSLPSTLYLGLPRKYVAALHLH----- 111

Qy 156 GTFQEFCCRFRMNDKLGYSWEFHVQEFW-----EHRMDSNVLFLLKYEDMHRD 202
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 112 -----WGQKSGPGGSEHQINSEATFAELHIVHYD 140

RESULT 4
US-10-511-814-1
; Sequence 1, Application US/10511814
; Publication No. US20060088472A1
; GENERAL INFORMATION:
; APPLICANT: McCance, Dennis
; APPLICANT: Westbrook, III, Thomas F.
; TITLE OF INVENTION: E7 REGULATION OF P21 (CIP1) THROUGH AKT
; FILE REFERENCE: 21108.0016U2
; CURRENT APPLICATION NUMBER: US/10/511,814
; CURRENT FILING DATE: 2004-10-19
; PRIOR FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: 60/374,245
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence;/Note =
US-10-511-814-1

Query Match          4.8%; Score 73.5; DB 6; Length 273;
Best Local Similarity 20.4%; Pred. No. 0.74;
Matches 50; Conservative 47; Mismatches 79; Indels 69; Gaps 14;

Qy 54 PKSGTSL-----QEVVLVSGADPPDEIGLNMIDEQLPVLEYPQPLGDIKELTSP 105
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 48 PDSFAALLSSNELGERGLVHVWAKALPGFRNLHVDQMAVIQYSWMGLMVA----- 102

Qy 106 RLKIKSHLPYRFLPSDLHNGDSKVIYMARNPDLVSVYYQFHR-----LRTMSYRGTFQEF 162
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 103 -----MGWR-----SFTVNSRMLYFA---PDLVFNRYMRKSRMYSOCVMRHLQSDF- 148

Qy 163 RRFMNDKLGYSWEFHV-QEFWEHRMDSNVLF-----LKYEDMHRDL-VTMVEQLARF 213
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 149 -----GWLQITPQEF--CMKALLFSIIPVDGLKNQKPFDELRMNVIKELDRI 195

Qy 214 LGVSCDKAQLAETHCHQLVDQCNAEALPVGRGRVGLWKDIFTVSNN-EKFDLVYKQK 272
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 196 --IAKRNKPTSCSRFFYQLTKLDSVQ--PIAR-----ELHQFTFDLLIKSH 239

Qy 273 MGKCD 277
| | | | |
Db 240 MVSVD 244
```

## RESULT 5

```
US-11-251-465-25
; Sequence 25, Application US/11251465
; Publication No. US20060094061A1
; GENERAL INFORMATION:
; APPLICANT: Brys, Reginald
; APPLICANT: Vandeginste, Nick
; APPLICANT: Tomme, Peter
; APPLICANT: Klaassen, Hubertus
; TITLE OF INVENTION: Molecular Targets And Compounds, And Methods To Identify The
; TITLE OF INVENTION: Same, Useful In The Treatment Of Joint Degenerative And
; TITLE OF INVENTION: Inflammatory Diseases
; FILE REFERENCE: P30,172-A USA
; CURRENT APPLICATION NUMBER: US/11/251,465
; CURRENT FILING DATE: 2005-10-14
; PRIOR APPLICATION NUMBER: 60/619,384
; PRIOR FILING DATE: 2004-10-15
; NUMBER OF SEQ ID NOS: 880
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-251-465-25

Query Match          4.7%; Score 72.5; DB 7; Length 365;
Best Local Similarity 19.5%; Pred. No. 1.5;
Matches 56; Conservative 50; Mismatches 116; Indels 65; Gaps 15;

Qy 11 TPGEFESKYFEFHGVRLPPFCRCGWBERIANFPVPSPDVWIVTPKSGTSLQEVVYLSQ 70
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 92 TPASSLRNFYDFYLVW--PFMQTDLQKIMGFSEKIQYLVY-----QMLUGLKYIHSA 144

Qy 71 GA-----DPDEIGLNMIDEQLPVLEY-----PQGLDIIKELTSPRLIKSHLPYR-- 115
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 145 GVVRDLKPGNLA-VNEDCELKILDFGLARHADAEIMGVYVTRYRAREVILSWHYNQT 203

Qy 116 -----FLPSDLHNGDSKVIYMARNPDLVSVYYQFHRSLRTMSYRGTFQEFCCRFRMND 168
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 204 VDIWSVGCIMAEMLTG--KTLFKGKDYLD-----QLTQILKVTGVPQT--EFVQK-LND 252

Qy 169 KLGYSWEFHVQEFWEHRMDSNVLFLLKYEDMHRDLVTMVEQLARFLGVSCDK--AQLEAL 226
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 253 KAA-KSYIQSLPQ--TPRKDFTLQFPRASPOAAL-----LEKMLELDVDRKLTAAQAL 303

Qy 227 TE-----HCHQLVDQCNAEALPVGRGRVGLWKDIFTVS 260
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 304 THPFFFRDPEETEAAQPPDDSLSEHEKLTVDWEKQHIYKEIVNFS 350
```

## RESULT 6

```
US-11-312-958-14
; Sequence 14, Application US/11312958
; Publication No. US20060100152A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Rosenfeld, Julie Beth
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 9949, 14230, 760, 62553,
; TITLE OF INVENTION: 12216, 17719, 41897, 47174, 33408, 10002, 16209, 314, 636,
; TITLE OF INVENTION: 27410, 33260, 619, 15985, 69112, 2158, 224, 615, 44373,
; TITLE OF INVENTION: 95431, 22245, 2387, 16658, 55054, 16314, 1613, 1675, 9569 OR
; TITLE OF INVENTION: 13424 MOLECULES
; FILE REFERENCE: MP102-027P1RNMNIM
; CURRENT APPLICATION NUMBER: US/11/312,958
; CURRENT FILING DATE: 2005-12-20
; PRIOR APPLICATION NUMBER: US/10/369,022
; PRIOR FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: US 60/360,495
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/370,121
```

```
/ PRIOR FILING DATE: 2002-04-04
/ PRIOR APPLICATION NUMBER: US 60/373,010
/ PRIOR FILING DATE: 2002-04-16
/ PRIOR APPLICATION NUMBER: US 60/373,908
/ PRIOR FILING DATE: 2002-04-19
/ PRIOR APPLICATION NUMBER: US 60/377,717
/ PRIOR FILING DATE: 2002-05-03
/ PRIOR APPLICATION NUMBER: US 60/379,949
/ PRIOR FILING DATE: 2002-05-13
/ PRIOR APPLICATION NUMBER: US 60/382,409
/ PRIOR FILING DATE: 2002-05-21
/ PRIOR APPLICATION NUMBER: US 60/385,280
/ PRIOR FILING DATE: 2002-06-03
/ PRIOR APPLICATION NUMBER: US 60/386,879
/ PRIOR FILING DATE: 2002-06-06
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 64
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 14
/ LENGTH: 367
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-196-749-466
```

```
Query Match      4.7%; Score 72.5; DB 7; Length 367;
Best Local Similarity 21.8%; Pred. No. 1.5; Mismatches 35; Indels 43; Gaps 9;
Matches 43; Conservative 35; Mismatches 35; Indels 43; Gaps 9;

Qy 54 PKSGTSL- QEVVYVLSQADPDEIGLNMIDQLPVLEYPQP-----GLDIIKELT 103
Db 107 PKLGTKRLPQALIVGVKGGTRAVLEFIRVHPDVRAAG-TEPHFPDRNYGRGLDWYRSL- 164

Qy 104 SPRLIKSHLPYRFLPSDL-----HNGDSKVIYMARPKDLVVSYYQFHRSLRTM 152
Db 165 MPTLESQITTEKTPSYFVTOEAPRRIFNMRDTKLIVVVRNPVTRAIISDY-----TQTL 219

Qy 153 SYR---GTFQFCRRFMNDKLG-----YGSWFHVOEFWEHRMDSNVFLFKYEDM 199
Db 220 SKKPDIPFEGGL--SFRNRTGLVDVSNNAIRIGMYVHLESWLQYFFLAQIHVFSGERL 277

Qy 200 HRDLVTMVEQLARFLGV 216
Db 278 ITDPAGEMGRVQDFLGI 294
```

```
RESULT 7
US-10-196-749-466
/ Sequence 466, Application US/10196749
/ Publication No. US20060094864A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430RIC340
/ CURRENT APPLICATION NUMBER: US/10/196,749
/ CURRENT FILING DATE: 2002-07-16
/ PRIOR APPLICATION NUMBER: 10/052586
/ PRIOR FILING DATE: 2002-01-15
/ PRIOR APPLICATION NUMBER: 60/059263
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/059266
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/062250
```

```
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/063120
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063121
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063486
/ PRIOR FILING DATE: 1997-10-21
/ PRIOR APPLICATION NUMBER: 60/063540
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063541
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063544
/ PRIOR FILING DATE: 1997-10-28
/ Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 466
/ LENGTH: 813
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-196-749-466
```

```
Query Match      4.7%; Score 72; DB 6; Length 813;
Best Local Similarity 21.3%; Pred. No. 5.8;
Matches 60; Conservative 37; Mismatches 85; Indels 100; Gaps 15;

Qy 42 PVPSDVIWITYPKSGTSLQEVVYVLSQADPDEIGLNMIDQLPVLEYPQPGLDIIKE 101
Db 39 PVTPH--WVLDGGQWRTVSLEEPV-----SKPD-MGLVALEAE-----GQELLLE 80

Qy 102 -----LTSPLRIKSHL-----PYRFLPSDLH-----NGDSKVI 129
Db 81 LEKNHRLLAGYIETHYGPDPQPVVLAPNHTDCHYQGRVGRPDSVNVLTCSGMSGLI 140

Qy 130 YMARNPDLVVSY-----OFHRSLTRMSYRGTFQEPFC-RRFNMNKLGYG 173
Db 141 TLSRN-----ASYVLRPWPGRGSKDFSTHEIFRMEQLLTWKGT-----CGHRDPGNKAGMT 191

Qy 174 SW-----PEHVQEFWE-HRMSNVFLFKYEDMHRDL-----VTMVQOLA 211
Db 192 SLPGGSGRGRREARRTKYKLYELIVADHTLFL---TTRHNLNHTKORLLELVANYVDQLL 249

Qy 212 RFLGVSCDKAQLAETEHCHQLVDQCCNAEALPVGRGRVGLW 253
Db 249 RTLDIQVALTGLEWTVTERDSRSRVTDANATLWAPLOWRGLW 290
```

```
RESULT 8
US-11-249-111-118
/ Sequence 118, Application US/11249111
/ Publication No. US20060099623A1
/ GENERAL INFORMATION:
/ APPLICANT: Glenn, Matthew
/ APPLICANT: Lubbers, Mark W
/ APPLICANT: Dekker, James
/ TITLE OF INVENTION: Polynucleotides and polypeptides isolated from Lactobacillus
/ FILE REFERENCE: 13353.1048uic2
/ CURRENT APPLICATION NUMBER: US/11/249,111
/ CURRENT FILING DATE: 2005-10-11
/ PRIOR APPLICATION NUMBER: 10/288,930
/ PRIOR FILING DATE: 2002-11-05
/ PRIOR APPLICATION NUMBER: 09/724,623
/ PRIOR FILING DATE: 2000-11-28
/ PRIOR APPLICATION NUMBER: 60/148,801
/ PRIOR FILING DATE: 1999-12-02
/ NUMBER OF SEQ ID NOS: 124
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 118
/ LENGTH: 259
/ TYPE: PRT
/ ORGANISM: Lactobacillus rhamnosus
US-11-249-111-118
```

Query Match 4.6%; Score 71; DB 7; Length 259;  
Best Local Similarity 24.6%; Pred. No. 1.3;  
Matches 32; Conservative 18; Mismatches 44; Indels 36; Gaps 7;  
  
QY 54 PKSGTSLQVYVLSQCADPD-----EIGLMNIDQLPVLVYPPQGLDII 99  
DB 66 PEEATELLWEAIFKTRIDQDPEAAWKAHQKLEKAALN-NEQFDQLHYMAFGTDLV 124  
QY 100 KELTSPRLIKSHL-----PYRFLPSDLHNGDSKVIYARNPKOLVVSYYOFHR-----SLR 150  
DB 125 VGLP-----KNHWEAGAFNP-----RGE-----FMANMPTBEVFTAPDFRRIDGTGVAATK 172  
QY 151 TMSYRGTFQE 160  
DB 173 PLSYGGNILE 182

RESULT 9

US-10-505-928-40  
; Sequence 40, Application US/10505928  
; Publication No. US20060088532A1  
; GENERAL INFORMATION:  
; APPLICANT: Ludwig Institute for Cancer Research et al.  
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES  
; FILE REFERENCE: 28967/39178  
; CURRENT APPLICATION NUMBER: US/10/505,928  
; CURRENT FILING DATE: 2004-08-27  
; PRIOR APPLICATION NUMBER: US 60/363,019  
; PRIOR FILING DATE: 2002-03-07  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: Patent in 3.2  
; SEQ ID NO 40  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-505-928-40

Query Match 4.5%; Score 69; DB 6; Length 1534;  
Best Local Similarity 22.1%; Pred. No. 33;  
Matches 55; Conservative 39; Mismatches 73; Indels 82; Gaps 14;  
  
QY 91 YPOQGLDIIKELTSRLIKSHLPYRFLPSDLHNGDS-KVIYMAR--NPKDL----- 138  
DB 1282 KYEPKLTPLSEI-SORLLK-----LYSDKPGSENKVMQDSKGNPKDLDSKYAYIQV 1333  
QY 139 --VVSYY-----QFHR-----LRTMSYRGTFQFCRRFMDKLGYS 174  
DB 1334 THVIPFDEKELQERKTEFERSHNIRRFMFEMPFTQTGKQGVVEQCKR--RTILTAIH 1391  
QY 175 WFEHVOE-----FWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHC 230  
DB 1392 CFPYVKRIPVMYQHHTDLNPIEVAIDEMSKK-VAELRQLCSSAEVDIMKLQLKLGYSVS 1450  
QY 231 HOLVQCCNAEALPVR-----GRVGLWKDIF-----TVSMNEKF--- 265  
DB 1451 VQV-----NAGPLAYARAFLODNTKRYPDNKKVLLKEVRFQVBEACQALAVNERLIKE 1505  
QY 266 -DLVYKQKM 273  
DB 1506 DQLEYQEM 1514

RESULT 10

US-11-311-778-18  
; Sequence 18, Application US/11311778  
; Publication No. US2006009598A1  
; GENERAL INFORMATION:  
; APPLICANT: Fischhoff, David A.  
; APPLICANT: Perlak, Frederick J.  
; TITLE OF INVENTION: SYNTHETIC PLANT GENES AND METHOD FOR PREPARATION  
; FILE REFERENCE: 38-21 (13553) C  
; CURRENT APPLICATION NUMBER: US/11/311,778  
; CURRENT FILING DATE: 2005-12-19

; PRIOR APPLICATION NUMBER: 08/434,105  
; PRIOR FILING DATE: 1995-05-03  
; PRIOR APPLICATION NUMBER: 07/959,506  
; PRIOR FILING DATE: 1992-10-09  
; PRIOR APPLICATION NUMBER: 07/476,661  
; PRIOR FILING DATE: 1990-02-12  
; PRIOR APPLICATION NUMBER: 07/315,355  
; PRIOR FILING DATE: 1989-02-24  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 18  
; LENGTH: 634  
; TYPE: PRT  
; ORGANISM: artificial  
; FEATURE:  
; OTHER INFORMATION: synthetic structural gene encoding P2 insecticidal protein  
US-11-311-778-18

Query Match 4.5%; Score 68.5; DB 7; Length 634;  
Best Local Similarity 19.4%; Pred. No. 9.6;  
Matches 45; Conservative 42; Mismatches 78; Indels 67; Gaps 13;  
  
QY 46 SDWIVTVPKSGTSLQEVV----YLVSQGADPDDEIGLMN-----IDEQ 85  
DB 76 SELWGIIFPGSGTNLMQDILRETEQFLNQLRLNTDTLARVNAELIGLOANIREFNQVDNF 135  
QY 86 LPVLEYPOPGLDIIKEL-TSPRLIKSHLP-----YRFL-----PSDLH----- 122  
DB 136 LNPTQNPVP-LSITSSVNTMQQLFLNRLPOFOIQYQLLLPLFAQAANMHLSTFIRDVIL 194  
QY 123 NGDSKVIYMA--RNPKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMDKLGYSW----- 175  
DB 195 NADEWGISAATLRTYRDYLRNYTRDYSNCINTYQTAFRGLNTR-LHDMLEFRYMFNLV 253  
QY 176 FEHVQEFWEHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLAALT 227  
DB 254 FEYV-SIWS-----LFKYQSL-----MVSSGANLYASGSGPQQTQST 290

RESULT 11

US-10-982-908-26  
; Sequence 26, Application US/10982908  
; Publication No. US20060099603A1  
; GENERAL INFORMATION:  
; APPLICANT: DAVI, FREDERIC  
; APPLICANT: DIGHERIO, GUILLAUME  
; APPLICANT: DUMAS, GERARD  
; APPLICANT: OPPEZO, PABLO  
; APPLICANT: SETTEGRANA, CATHERINE  
; APPLICANT: VASCONCELOS PINHEIRO, YURI  
; TITLE OF INVENTION: METHOD OF DIAGNOSIS/PROGNOSIS OF HUMAN CHRONIC LYMPHOCTIC  
; FILE REFERENCE: 259358USO  
; CURRENT APPLICATION NUMBER: US/10/982,908  
; CURRENT FILING DATE: 2004-11-08  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: Patent in version 3.3  
; SEQ ID NO 26  
; LENGTH: 820  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-982-908-26

Query Match 4.5%; Score 68.5; DB 6; Length 820;  
Best Local Similarity 19.2%; Pred. No. 14;  
Matches 71; Conservative 42; Mismatches 81; Indels 175; Gaps 21;  
  
QY 31 CRGKMEIEIANPVPSPDV---WI-----VTPKSGTSLQEVVLSQSG 71  
DB 173 CRMEFEIDNSTQKQSSVYGVGWHIFRIVEIVVVDNVLVIRYERNDSKLED-LYVIVNI 231  
QY 72 ADP--DEIGLMNIDEQLPVLEYPOPGLDIKELTSPRI-----KS-HLPYRFLPSDLH 122

```
Db 232 VDSILDVIG-----VKVLLF---GLEI---WTNKNLIVDDVRKSVHLYCKWKSENI- 277
Qy 123 NGDSKVIYARNPKDLVVSYVQFHRSLRTMSYRGTFQBFQR-----RPMNDKLG--- 171
Db 278 -----TPRMQHD--TSHLFTTLGLRGLSGIGAFRGMCPTPHRSCAIVTFPMNKTGLTFS 327
Qy 172 -----YGSFHFVQBEWEHRMD 188
Db 328 IAVAHHLGHNLGMNHDEDTCSQPCRMHGNPPITKFSNCSYG-----DFWEYTYE 380
Qy 189 SNVLFLKYEDMHRDLVTMVEQLARFL---GVSCDKAQLEALTEHCHQLVDOCC----- 238
Db 381 RTKCLL--ETVHTKDI-FNVKRCGVVBECEDCGPL-----KHCAK--DFCCLSNCTLT 432
Qy 239 -----NARALPVGRRVGLWKDIFTVSMNEKFDLVYKQKMGKCDL----- 278
Db 433 DGSTCAFGLCCKDCKFLPSGK-----VCRKEVNECDLPENCNGTS 472
Qy 279 ---TFDFYL 284
Db 473 HKCPDPFYV 481

RESULT 12
US-11-242-505A-33
; Sequence 33, Application US/11242505A
; Publication No. US20060099656A1
; GENERAL INFORMATION:
; APPLICANT: Carroll, Joseph M.
; TITLE OF INVENTION: Methods and Compositions for Treating
; TITLE OF INVENTION: Hematological Disorders Using 232, 2059, 10630, 12848, 13875,
; FILE OF INVENTION: 14395, 14618, 17692, 58874, 252, 304, 1980, 14717, 9941, 19310,
; FILE REFERENCE: MPI2001-288PRCPTOMNIM
; CURRENT APPLICATION NUMBER: US/11/242.505A
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US 10/290,078
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: US 60/347,949
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 10/320,351
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,606
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-242-505A-33

Query Match 4.4%; Score 67.5; DB 7; Length 424;
Best Local Similarity 30.3%; Pred. No. 6.7;
Matches 27; Conservative 13; Mismatches 40; Indels 9; Gaps 4;

Qy 38 IANFVRSDWIVTPKSGTSL--QEVVYL--VSCGADDPDEIGLMNIDQL-----PVLE 90
Db 83 VALFQVLPQLMDITYRFQGPDLRCRAVKYQLQVLSMFASYMLLANTLDRVLAVCHPLRS 142
Qy 91 YPQPGLDIIKELTSPLRIKS--HLPVREL 117
Db 143 LQPGQGVSTYLLIAAPWLLAAIFSLQVFI 171

RESULT 13
US-10-505-928-225
; Sequence 225, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178

; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 225
; LENGTH: 1056
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-225

Query Match 4.4%; Score 67.5; DB 6; Length 1056;
Best Local Similarity 18.2%; Pred. No. 27;
Matches 52; Conservative 43; Mismatches 92; Indels 99; Gaps 11;

Qy 35 MEEIANFVRPSDVIWIVTPKSGTSL--QEVVYL--VSCGADDPDEIGLMNIDQLPVLEYQPQ 94
Db 303 VERTPHVPYRESKLRILODLSLGRTRTSIATISPAS-----LNLEETLSTLEYAHR 355
Qy 95 GLDII-----KELTSPLRIKSHL--PYRFLPSDLHNGDSKVIYARNPKDLVVSYYQFHR 147
Db 356 AKNILNKPENQKLTKKALIKETTEIERLKRDL-----AAAREKNGVYIISEENF-- 405
Qy 148 SLRTMSYRGTFQE-----FCRRFMNDKLGYSW 175
Db 406 --RVMSGKLTVOBEQIVLEIKIGAVEEELNRVTLEFMDNKNELDQCKSLQNTK----- 458
Qy 176 FEHVQEFWEHRMDSNVLFKYE-----DMHRDLVTMVBQLARFL--GVSCDK 220
Db 459 -QELTTQKHLQETKQLVKEEYITSALESTEELKHAASKLLNTVETTTKDVSGLSK 517
Qy 221 AQLBALTEHCHQLVDQCCNARALPVGRRVGLWKDIFTVSMNEKFD 266
Db 518 DRKKAVDQH-----NABA-----QDIFGKNLSLNF 543

RESULT 14
US-10-511-937-2407
; Sequence 2407, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2407
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2407

Query Match 4.2%; Score 65; DB 6; Length 587;
Best Local Similarity 24.7%; Pred. No. 21;
Matches 42; Conservative 24; Mismatches 56; Indels 48; Gaps 9;
```

```

; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Defrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bowe, Caryne
; TITLE OF INVENTION: ERYTHROPOIETIN: REMODELING AND
; TITLE OF INVENTION: GLYCOCONJUGATION OF ERYTHROPOIETIN
; FILE REFERENCE: 040853-01-5083-US02
; CURRENT APPLICATION NUMBER: US/11/183,218
; CURRENT FILING DATE: 2005-07-15
; PRIOR APPLICATION NUMBER: US 10/410,945
; PRIOR FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: PCT/US02/32263
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: US 60/334,301
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 60/334,233
; PRIOR FILING DATE: 2001-11-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-183-218-42

Query Match 4.2%; Score 64; DB 7; Length 226;
Best Local Similarity 25.0%; Pred. No. 6.2;
Matches 36; Conservative 17; Mismatches 41; Indels 50; Gaps 7;

QY 36 ERIANFVRPSDVWIVYTPKSGTSLQEVVYLVSQGADPDPEIGLMNTDEQLP----- 87
DB 81 EDIADYVCOQSHSWPPTFF-GSGTNLVKEYVKLEESGG-----GLVQPGGKMLSCVASGF 134
QY 88 -----VLEYPQGLDITIKELTSPRLTKSHLPY-----RFLPS----- 119
DB 135 IFSNHNWNVVRSPEKGLWEVAIRS-KSNSATHVAESVKGRTISRDRDSKSAVYLQMT 193
QY 120 DLHNGDSKVIYMARNPKDLVWSY 143
DB 194 DLRTEDTGVVYCSRN-----YY 210

RESULT 17
US-10-196-749-270
; Sequence 270, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.

```

```

RESULT 18
US-10-984-694-3
; Sequence 3, Application US/10984694
; Publication No. US20060099679A1
; GENERAL INFORMATION:
; APPLICANT: TSJEN, Roger Y.
; APPLICANT: WANG, Lei
; TITLE OF INVENTION: METHODS FOR ENGINEERING POLYPEPTIDE
; TITLE OF INVENTION: VARIANTS VIA SOMATIC HYPERMUTATION AND POLYPEPTIDES MADE
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 39754-0976A US
; CURRENT APPLICATION NUMBER: US/10/984,694
; CURRENT FILING DATE: 2004-11-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fluorescent protein (mPlum)
US-10-984-694-3

Query Match          4.2%; Score 63.5; DB 6; Length 226;
Best Local Similarity 21.7%; Pred. No. 7;

```



```
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0207904.4
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0207900.2
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0207901.0
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0227734.1
; PRIOR FILING DATE: 2002-11-28
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-455-50

Query Match
Best Local Similarity 4.1%; Score 62.5; DB 6; Length 933;
Matches 47; Conservative 28; Mismatches 62; Indels 107; Gaps 10;

QY 26 RLPPFCRGMBEIA-----NFPVRSDVWIVTYPKSGTSLQGVVYLVVSQGADPDEIGLM 80
Db 420 RYEGFCVDMKELAEILRFNYKIRLVGDGVYGPVPEANGTWTGMVGLIARKADLAVAGLT 479
QY 81 NIDEQLPVLEYPOP-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYARNPK 136
Db 480 ITAEREKVIDFSKPFMTLGISIL-----YR-----IHMGRKP- 511
QY 137 DLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVFQEFWEHRMDSNVLFKY 196
Db 512 ----GYFSF-----LDPFSPGW-----LFM-- 528
QY 197 EDMHRDLVTMVEQLARFLGVSCDKAQLAEALTEH-----CHQLVDQCCNAEAL- 243
Db 529 -----LLAYLAVSCVLFVARLTPYEWYSPHPCAQGRCNLLVNOYSLGSLW 575
QY 244 -PVG 246
Db 576 FPVG 579

RESULT 21
US-10-511-455-41
; Sequence 41, Application US/10511455
; Publication No. US2006008835A1
; GENERAL INFORMATION:
; APPLICANT: Pickard, Benjamin Simon
; APPLICANT: Blackwood, Douglas
; APPLICANT: Porteous, David
; APPLICANT: Muir, Walter John
; APPLICANT: Mors, Ole
; APPLICANT: Ewald, Henrik Lykke
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES
; FILE REFERENCE: 9013.63
; CURRENT APPLICATION NUMBER: US/10/511,455
; PRIOR FILING DATE: 2004-10-05
; PRIOR APPLICATION NUMBER: PCT/GB03/001543
; PRIOR FILING DATE: 2003-04-07
; PRIOR APPLICATION NUMBER: GB0207902.8
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0207904.4
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0207900.2
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0207901.0
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0227734.1
; PRIOR FILING DATE: 2002-11-28
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 956
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-455-41

Query Match
Best Local Similarity 4.1%; Score 62.5; DB 6; Length 956;
Matches 47; Conservative 28; Mismatches 62; Indels 107; Gaps 10;

QY 26 RLPPFCRGMBEIA-----NFPVRSDVWIVTYPKSGTSLQGVVYLVVSQGADPDEIGLM 80
Db 443 RYEGFCVDMKELAEILRFNYKIRLVGDGVYGPVPEANGTWTGMVGLIARKADLAVAGLT 502
QY 81 NIDEQLPVLEYPOP-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYARNPK 136
Db 503 ITAEREKVIDFSKPFMTLGISIL-----YR-----IHMGRKP- 534
QY 137 DLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFHVFQEFWEHRMDSNVLFKY 196
Db 535 ----GYFSF-----LDPFSPGW-----LFM-- 551
QY 197 EDMHRDLVTMVEQLARFLGVSCDKAQLAEALTEH-----CHQLVDQCCNAEAL- 243
Db 552 -----LLAYLAVSCVLFVARLTPYEWYSPHPCAQGRCNLLVNOYSLGSLW 598
QY 244 -PVG 246
Db 599 FPVG 602

RESULT 22
US-11-283-522-34
; Sequence 34, Application US/11283522
; Publication No. US2006008914A1
; GENERAL INFORMATION:
; APPLICANT: Pan, James
; APPLICANT: Adams, Sean
; TITLE OF INVENTION: UCP5
; FILE REFERENCE: P1663R2
; CURRENT APPLICATION NUMBER: US/11/283,522
; CURRENT FILING DATE: 2005-11-18
; PRIOR APPLICATION NUMBER: US/10/270,861
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US/09/433,622
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: US 60/110,286
; PRIOR FILING DATE: 1998-11-30
; PRIOR APPLICATION NUMBER: US 60/129,583
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: US 60/143,886
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 34
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-283-522-34

Query Match
Best Local Similarity 4.1%; Score 62; DB 7; Length 309;
Matches 30; Conservative 18; Mismatches 43; Indels 28; Gaps 6;

QY 97 DIIKE-LTSPRLIKSHLPYRFLPSDLHNGDSKVIYARNPKDLVVSY------QF----- 145
Db 198 DLIKDALIKANLMTDDLPCFTSA---FGAGFCTTIVIASPDVVVKTRYMNSALGOYSAG 254
QY 146 HRSRLTMSYRGTFQFCRRFMNDKLGYSWFHVFQEFWEHRMDSNVLFKYEDMHRDLV 204
Db 255 HCAITMLQKEGP-RAFYKGFMPSFLRLGSW-----NVVMFVTEQLKRALM 299

RESULT 23
US-10-505-928-642
; Sequence 642, Application US/10505928
```



Qy	69	SQAGDPDEIGLWNIDEQLPVLEYPGQGLDIKELTSPR--LIKSHLPVRYFLPSD--LHNG	124
Db	124	-----KLRTGNPDSGVPVMOKTMGTWEASSERMYPBDGALKGEMKMRLLKLDGGHYDA	176
Qy	125	DSKVITYMARNPKDLVVSY	142
Db	177	EVKTTYMAKKPVOLPGAY	194

```

RESULT 27
US-10-505-928-233
; Sequence 233, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 233
; LENGTH: 512
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-505-928-233

```

```

RESULT 28
US-10-511-937-2512
; Sequence 2512, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; ; ; ; ;
; FILE REFERENCE: 50661200104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2

; ORGANISM: HOMO SAPIENS
US-11-211-917-40

Query Match          4.0%; Score 61; DB 7; Length 239;
Best Local Similarity 22.8%; Pred. No. 14;
Matches 39; Conservative 25; Mismatches 57; Indels 50; Gaps 11

Qy   1 MAESEAEETPTGCFESKYFEFHGVRLPPFCRGKMEIEANFVRPSDVMIVTVPKSGTSL 60
      | .| ||||| : : : : :
Db   24 MTQSPLSLPVTGPEDAS-----ISCRSSQSVLYSGVNYLD-WYLOKPGGSPQL 71

Qy   61 LQEVVYLVSQGAD--PDE-----IGLMNIDEQ-----LPVLVEYP---OPG-- 95
      | :||| ||| : : : : :
Db   72 L----IYLGSNRASGVDFRFGSGSGTDFTLKTSRVEAEDVGYYCQMVLQTFFTFPGGYK 128

Qy   96 LDITKELTSPRLIKSHLPIRYFLPSD--LHNGDSKIYIMARN--PKDLVVSY 142
      :|| : :||| : : : : :
Db   129 VDIKRTVAAPSV-----FIFFPSDQLKSAGTSVVCLLNFFVPBREAKVQW 173

RESULT 30
US-10-511-937-2584
; Sequence 2584, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay

```



Publication No. US20060088906A1  
GENERAL INFORMATION:  
APPLICANT: Necse Technologies, Inc.  
APPLICANT: Defrees, Shawn  
APPLICANT: Zopf, David  
APPLICANT: Bayer, Robert  
APPLICANT: Hakes, David  
APPLICANT: Chen, Xi  
APPLICANT: Bowe, Caryne  
TITLE OF INVENTION: ERYTHROPOIETIN: REMODELING AND  
FILE REFERENCE: 040853-01-5083-US02  
CURRENT APPLICATION NUMBER: US/11/183,218  
CURRENT FILING DATE: 2005-07-15  
PRIOR APPLICATION NUMBER: US 10/410,945  
PRIOR FILING DATE: 2003-04-09  
PRIOR APPLICATION NUMBER: PCT/US02/32263  
PRIOR FILING DATE: 2002-10-09  
PRIOR APPLICATION NUMBER: US 60/407,527  
PRIOR FILING DATE: 2002-08-28  
PRIOR APPLICATION NUMBER: US 60/404,249  
PRIOR FILING DATE: 2002-08-16  
PRIOR APPLICATION NUMBER: US 60/396,594  
PRIOR FILING DATE: 2002-07-17  
PRIOR APPLICATION NUMBER: US 60/391,777  
PRIOR FILING DATE: 2002-06-25  
PRIOR APPLICATION NUMBER: US 60/387,292  
PRIOR FILING DATE: 2002-06-07  
PRIOR APPLICATION NUMBER: US 60/344,692  
PRIOR FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: US 60/334,301  
PRIOR FILING DATE: 2001-11-28  
PRIOR APPLICATION NUMBER: US 60/334,233  
PRIOR FILING DATE: 2001-11-28  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 75  
SOFTWARE: Patent in version 3.2  
SEQ ID NO 30  
LENGTH: 2351  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-11-183-218-30

Query Match 3.9%; Score 60; DB 7; Length 2351;  
Best Local Similarity 20.0%; Pred. No. 6.1e+02;  
Matches 51; Conservative 39; Mismatches 85; Indels 80; Gaps 14;  
QY 53 YPKSGTSLQEVVYLVSGADPDEIGLMNIDEQLPVLEYQPGLDIIKELTSPRLIK--S 110  
Db 473 YGEVGD TLL---IIFKNQASRPYNI-----YPH-GITDVRPLYSRRLPKGVK 515  
QY 111 HL-PYRFLPSDL-----HNGDSKVIYMARNPKDLVVSYYQFHRSLRTWS----- 153  
Db 516 HLKDPFLPGEIFKYKWTVTVEDGPTK-----SDPCLTRYYSFVNMRDLASGLIGPL 570  
QY 154 ---YRGTFQFCRRFMNDKLGYSWPEHVQEFWEHRMDSNVLFLKYEDMHRDLVTMVEQL 210  
Db 571 LICYKESVDQRGNQIMSDK-----RNVILFSVFDENRSWY--LTENI 610  
QY 211 ARFL-----GVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRGVGLWKDIFTVSMNEKPD 266  
Db 611 QRFLPNPAGVQLEDPEFOA-SNIMHSINGYVDFSLQLSVCLHEVAYW---YILSIGAQTD 666  
QY 267 LV-----YKQKM 273  
Db 667 FLVSFFSGYTFKHKM 681

RESULT 34  
US-11-280-757-35  
Sequence 35, Application US/11280757  
Publication No. US20060099685A1  
GENERAL INFORMATION:

APPLICANT: Crucell Holland B.V.  
APPLICANT: Bout, Abraham  
APPLICANT: Opstelten, Dirk-Jan  
APPLICANT: Yallop, Christopher  
TITLE OF INVENTION: RECOMBINANT EXPRESSION OF FACTOR VIII IN HUMAN CELLS  
FILE REFERENCE: 0034 D US P00 CIP  
CURRENT APPLICATION NUMBER: US/11/280,757  
CURRENT FILING DATE: 2005-11-15  
PRIOR APPLICATION NUMBER: US 10/234,007  
PRIOR FILING DATE: 2002-09-03  
PRIOR APPLICATION NUMBER: US 09/549,463  
PRIOR FILING DATE: 2000-04-15  
PRIOR APPLICATION NUMBER: 60/129,452  
PRIOR FILING DATE: 1999-04-15  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: Patent in version 3.3  
SEQ ID NO 35  
LENGTH: 2351  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-11-280-757-35  
Query Match 3.9%; Score 60; DB 7; Length 2351;  
Best Local Similarity 20.0%; Pred. No. 6.1e+02;  
Matches 51; Conservative 39; Mismatches 85; Indels 80; Gaps 14;  
QY 53 YPKSGTSLQEVVYLVSGADPDEIGLMNIDEQLPVLEYQPGLDIIKELTSPRLIK--S 110  
Db 473 YGEVGD TLL---IIFKNQASRPYNI-----YPH-GITDVRPLYSRRLPKGVK 515  
QY 111 HL-PYRFLPSDL-----HNGDSKVIYMARNPKDLVVSYYQFHRSLRTWS----- 153  
Db 516 HLKDPFLPGEIFKYKWTVTVEDGPTK-----SDPCLTRYYSFVNMRDLASGLIGPL 570  
QY 154 ---YRGTFQFCRRFMNDKLGYSWPEHVQEFWEHRMDSNVLFLKYEDMHRDLVTMVEQL 210  
Db 571 LICYKESVDQRGNQIMSDK-----RNVILFSVFDENRSWY--LTENI 610  
QY 211 ARFL-----GVSCDKAQLEALTEHCHQLVDQCCNAEALPVGRGVGLWKDIFTVSMNEKPD 266  
Db 611 QRFLPNPAGVQLEDPEFOA-SNIMHSINGYVDFSLQLSVCLHEVAYW---YILSIGAQTD 666  
QY 267 LV-----YKQKM 273  
Db 667 FLVSFFSGYTFKHKM 681  
RESULT 35  
US-11-101-316-140  
Sequence 140, Application US/11101316  
Publication No. US20060099657A1  
GENERAL INFORMATION:  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID  
FILE REFERENCE: P3230R1C17C1  
CURRENT APPLICATION NUMBER: US/11/101,316  
CURRENT FILING DATE: 2005-04-06  
PRIOR APPLICATION NUMBER: 10/063526  
PRIOR FILING DATE: 2002-05-03  
PRIOR APPLICATION NUMBER: 10/006867  
PRIOR FILING DATE: 2001-12-06  
PRIOR APPLICATION NUMBER: PCT/US00/23328  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: 09/380137  
PRIOR FILING DATE: 1999-08-25  
PRIOR APPLICATION NUMBER: PCT/US99/12252  
PRIOR FILING DATE: 1999-06-02  
PRIOR APPLICATION NUMBER: 60/087759

;; PRIOR FILING DATE: 1998-06-02  
;; NUMBER OF SEQ ID NOS: 170  
;; SEQ ID NO 140  
;; LENGTH: 310  
;; TYPE: PRT  
;; ORGANISM: Homo Sapien  
US-11-101-316-140

Query Match 3.9%; Score 59.5; DB 7; Length 310;  
Best Local Similarity 20.5%; Pred. No. 32;  
Matches 46; Conservative 26; Mismatches 77; Indels 75; Gaps 11;

Qy 2 AESAEATPS--TPGEFESKYFHFHGVRLPPFCRGKMEIEANFPVRPSDVWIVTPYKSGTSL 60  
Db AEHGAEPAPPSGASS-----NWTLPPLPFSKVIVLIDALRDFV-----FGSKGVKF 89

Qy 61 LQEVVYLSQGAADPEIGLMNIDEQLPVLEYPQ-----PG-LDIKELTSPRLIKS 110  
Db 90 MPYTTLVEKGASHSFVA---EAKPPTVTMPRIKALMTGSLPGFVDVIRNLNSPALL-- 143

Qy 111 HLPYRFLPSDLHNGDSKVIYARNPKOLUVSY-----YQFHSRLTMSYRGTFQEF- 161  
Db 144 -----EDSVIROAKAAGKRIVFYGDETWVKLFPKH-----FVEYDGTTSFFV 185

Qy 162 -----CRRFMNDKLGYSWFHVOEFWEHRMDSNVLEFKY 196  
Db 186 SDYTEVDNNVTRHLDKVLKRGDW-----DILILHY 215

RESULT 36  
US-10-196-749-404  
; Sequence 404, Application US/10196749  
; Publication No. US20060094864A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C340  
; CURRENT APPLICATION NUMBER: US/10/196,749  
; CURRENT FILING DATE: 2002-07-16  
; PRIOR APPLICATION NUMBER: 10/052586  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059266  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063120  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063121  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063486  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/063540  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063541  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063544  
; PRIOR FILING DATE: 1997-10-28  
; Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 404

;; LENGTH: 347  
;; TYPE: PRT  
;; ORGANISM: Homo Sapien  
US-10-196-749-404

Query Match 3.9%; Score 59.5; DB 6; Length 347;  
Best Local Similarity 31.4%; Pred. No. 37;  
Matches 16; Conservative 10; Mismatches 12; Indels 13; Gaps 4;

Qy 27 LPPF-CRGKMEIEANFPVRPSDVWIV--TYPKSGTSLQEVVYLSQGADP 74  
Db 295 IPPPQCRHRCQSA-MBIEPGDIGYDVTTHWK-----VYVIARGVQP 335

RESULT 37  
US-11-101-316-124  
; Sequence 124, Application US/11101316  
; Publication No. US20060099657A1  
; GENERAL INFORMATION:  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID  
; FILE REFERENCE: P3230R1C17C1  
; CURRENT APPLICATION NUMBER: US/11/101,316  
; CURRENT FILING DATE: 2005-04-06  
; PRIOR APPLICATION NUMBER: 10/063526  
; PRIOR FILING DATE: 2002-05-03  
; PRIOR APPLICATION NUMBER: 10/006867  
; PRIOR FILING DATE: 2001-12-06  
; PRIOR APPLICATION NUMBER: PCT/US00/23328  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 09/380137  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: PCT/US99/12252  
; PRIOR FILING DATE: 1999-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; NUMBER OF SEQ ID NOS: 170  
; SEQ ID NO 124  
; LENGTH: 513  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-11-101-316-124

Query Match 3.9%; Score 59.5; DB 7; Length 513;  
Best Local Similarity 24.4%; Pred. No. 68;  
Matches 44; Conservative 19; Mismatches 50; Indels 67; Gaps 11;

Qy 31 CRGKMEIEANFPVRPSDVWIVTPYKSGTSLQEVVYLSQGADPEIGLMNIDQLPVLE 90  
Db 40 CEGKMK-----VYCES--QKLEIPSSISAGC-----LGLSLRYNSLQKLK 77

Qy 91 YPO-PGLDIIKELTSPRLIKSHLPYRFLPSPDLHG-----DSKVIYMARNP----- 135  
Db 78 YNQPKGLN---OLTWLYLDHNHISN--IDENAFNGIRRLKELILSSNRISFLNNTFRPV 132

Qy 136 ---KDLVVSYQFHF-----RSLRTMSYRGTFQEFRRFMNDKLG 172  
Db 133 TNLNRLDLSYNQLHSLGSEQFRGLRKULLSLHLSRNSLRTIPVR-IFQD-CRNLELLDLY 190

RESULT 38  
US-11-252-080-2  
; Sequence 2, Application US/11252080  
; Publication No. US20060090217A1  
; GENERAL INFORMATION:  
; APPLICANT: Texas A&M University  
; TITLE OF INVENTION: ISOLATION OF SUGARCANE PROTEINS INVOLVED IN POSTTRANSCRIPTIONAL  
; TITLE OF INVENTION: SILENCING AND METHODS OF USE







Matches	63;	Conservative	45;	Mismatches	116;	Indels	93;	Gaps	16
Qy	29	PFCRGKMEEIAN----	FVVRPSDV-----	WIVTPK-----	SGETLLQEWYLV	68			
Dd	161	PMCSYQMERMTTTRIPGKD	TDLVLQLHSDSRHVAVYHKGRFFKLWLYEGARLLK-----	214					
Qy	69	SQGADPDEIGLM----	NDEQLPVLEYPOPGLDIIKELTS-	PRLIKSHLPYRFLPSDLNG	124				
Dd	215	-----	PDLEMQFORILDDPS- -	POPGBEKLAALTAGGRVEWAQAQAFSSGKNKA	265				
Qy	125	DSKVINYMAR-----	NPKD-LVVSYY-----	-QFHRSLRTMSYRGTFQ	159				
Dd	266	ALEAIERAAPFVALDESYSYDP	EDEASLSLYGKALLHCNCYNRFDKSFTLIS-----	319					
Qy	160	EFCRRFNNDKLGYSWFPH-----	VQSFWEHRMDSNVLFKYEDMRDLVTMVVEQLAR	212					
Dd	320	-----	FKNQQLGLNA--EHAWADAPIIGHLWEFVLGTDSFHLGYTTETGHCLGKPNPALAP	372					
Qy	213	FLGVSCD-KAQLEALTEHQOL-----	VDQCNAEALPVGRGRVGLWKIDIFTVSMNEK	264					
Dd	373	PTRLQWDIPKOCQAVIESSYQAKALADDVELYC-FQFLPFKGKLIKKICRTSPDAFVOIA	431						
Qy	265	FDLVYVKOMGKCIDLTFD	281						
Dd	432	LQLAHFRDRGRKFCLTYE	448						
<b>RESULT 45</b>									
US-10-505-928-644									
; Sequence 644, Application US/10505928									
; Publication No. US20060088532A1									
; GENERAL INFORMATION:									
; APPLICANT: Ludwig Institute for Cancer Research et al.									
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES									
; FILE REFERENCE: 28967/39178									
; CURRENT APPLICATION NUMBER: US/10/505,928									
; PRIORITY FILING DATE: 2004-08-27									
; PRIOR APPLICATION NUMBER: US 60/363,019									
; NUMBER OF SEQ ID NOS: 866									
; SOFTWARE: PatentIn 3.2									
; SEQ ID NO 644									
; LENGTH: 727									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-10-505-928-644									

Db 420 GALSADGYARSEDYVDIVQRRVGWVNNPYISNIYL 456

RESULT 46

US-10-511-937-2606

; Sequence 2606, Application US/10511937

; Publication No. US20060088836A1

; GENERAL INFORMATION:

; APPLICANT: EXPRESSION DIAGNOSTICS, INC.

; APPLICANT: Wohlgenuth, Jay

; APPLICANT: Fry, Kirk

; APPLICANT: Woodward, Robert

; APPLICANT: Ly, Ngoc

; APPLICANT: Prentice, James

; APPLICANT: Morris, MacDonald

; APPLICANT: Rosenberg, Steven

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING

; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION

; FILE REFERENCE: 506612000104

; CURRENT APPLICATION NUMBER: US/10/511,937

; CURRENT FILING DATE: 2004-10-19

; PRIOR APPLICATION NUMBER: PCT/US2003/012946

; PRIOR FILING DATE: 2003-04-24

; PRIOR APPLICATION NUMBER: US 10/131,831

; PRIOR FILING DATE: 2002-04-24

; PRIOR APPLICATION NUMBER: US 10/325,899

; PRIOR FILING DATE: 2002-12-20

; NUMBER OF SEQ ID NOS: 3117

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 2606

; LENGTH: 727

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-511-937-2606

Query Match 3.8%; Score 58.5; DB 6; Length 7

Best Local Similarity 21.4%; Pred. No. 1.5e+02;

Matches 72; Conservative 48; Mismatches 116; Indels

Qy 1 MAESEAETPTGFEESKYP-----EFHCVRLLPPCR-----GKWE

Db 168 VAEWGGDSQDLFTKIFLDPKEKQINITLDHRCRIFQNLDGALD

Qy 45 PSDVWIVTYP-----KSGTSLLQLWVYL-----VSQCADPDEIGLA

Db 228 ARNLAYDTLPVLHGNGPTKQL-LNLGVNYPFWTFETGCTVCDSEGL

Qy 88 V-----LEVPGQLDIIKELTSPRIKSHLPVRFLPSPDLHNGDSKVI

Db 287 TVLVGVFIEQTPFVSLFFQ-----RLLLRHPQKHMRLFIHN-----

Qy 142 YYQFHRSLRTMSYRGTFQBFRCFRFMDKLGYSWFSHFVQEFW-EHRMD

Db 325 HEQHHKA-----QVEEFLAQ-----HGSEQSVKLVGPEVRMA

Qy 201 RDLVTMVQELARFLGVSCDKAQALEALTE--HCHQLVDQCNAEA-LPVC

Db 366 ADLCRDRSCITYFSDAD-----VALTEPNSLELLIQKNKVIAPLMT

Qy 255 -----DIFTVSMNEKFDLVYKQKMGKCDJTF--DFYL 284

Db 420 GALSADGYARSEDYVDIVQRRVGWVNNPYISNIYL 456

RESULT 47

US-10-505-928-304

; Sequence 304, Application US/10505928

; Publication No. US20060088532A1

; GENERAL INFORMATION:

; APPLICANT: Ludwig Institute for Cancer Research et al.

; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES

; FILE REFERENCE: 28967/39178

; CURRENT APPLICATION NUMBER: US/10/505,928

```

RESULT 47
US-10-505-928-304
; Sequence 304, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505.928

```

```
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 304
; LENGTH: 763
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-304
```

```
Query Match 3.8%; Score 58.5; DB 6; Length 763;
Best Local Similarity 31.4%; Pred. No. 1.6e+02;
Matches 27; Conservative 11; Mismatches 33; Indels 15; Gaps 5;

QY 44 RPSDVWIVTPKSGTSLQEVVYLVSQGAD---PDEIGLMNID----BQLPVLVYPQGL 96
Db 619 RTPESWRLTPPAKVGGLDFSPVQ-TSQGADPLPDPLGLMDLSTTPLQSAAPLESFQ--- 674

QY 97 DIKELTSPRLIKSHLPY-RFLPSDL 121
Db 675 ---RLLSPELDLISVPRGNSPSPDI 697
```

```
RESULT 48
US-10-505-928-449
; Sequence 449, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
```

```
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 449
; LENGTH: 3396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-449
```

```
Query Match 3.8%; Score 58.5; DB 6; Length 3396;
Best Local Similarity 21.7%; Pred. No. 1.5e+03;
Matches 40; Conservative 27; Mismatches 66; Indels 51; Gaps 10;

QY 10 STPGFESKYFEFHGVRLPFCRCRKMEETANFPVRPSDVWIVTPKSGTSL---QEVV 65
Db 2515 STDGSGQDRFREFFEDSTLKNRK-----KPTENIIDDKDKDLITITESTI 2563

QY 66 YLVSQGADPDEIGLMNIDBQLPVLVYPQGLDIIKELTSPRLIKSHLPYRFLPSDLH--- 122
Db 2564 LEILPELTSRDKNTIIDIDHTKEVYE-----DILGQMTD---IDTEVFSE--PHDSNDES 2612

QY 123 NGDS---KVIYMA-----RNPXDLVVSYYQFIRSLTMSYRGTFQFCRRFMD 168
Db 2613 NDDSTQVOBIYEAANVLSITETFECSADVLASYTQATHD-ESMTYED-----RSQLD 2664

QY 169 KLGY 172
Db 2665 HMGF 2668
```

```
RESULT 49
US-10-196-749-496
; Sequence 496, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
```

```
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 496
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-496
```

```
Query Match 3.8%; Score 58; DB 6; Length 458;
Best Local Similarity 27.1%; Pred. No. 84;
Matches 32; Conservative 18; Mismatches 44; Indels 24; Gaps 7;

QY 22 FHGVRLLPFCRCRKMEETANFPVRPSDVWIVTPKSGTSLQEVVYLVSQGADPDEIGLMN 81
Db 337 FDGVYQPL-RGQFVVEASYP--GQDRWLRYCASGLYILT---LLHEGYGFSE----- 384

QY 82 IDEQLPVLEYQPQ--GLDI-----IKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYM 131
Db 385 --ETWPSLEFRKQAGGVDIGMTGLYMLNLTG--MIPADAPQWRAESYGVVWVAKVVF 438
```

```
RESULT 50
US-10-473-691B-2
; Sequence 2, Application US/10473691B
; Publication No. US20060099202A1
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles A.
; APPLICANT: Soltis, Daniel A.
; TITLE OF INVENTION: IMMUNOGLOBULIN CONSTRUCT CONTAINING TUMOR SPECIFIC p53BP2 SEQUENC
; FILE REFERENCE: 02755/100K313-US1
; CURRENT APPLICATION NUMBER: US/10/473,691B
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: PCT/US02/10224
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: US 60/280,733
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
```



```
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 416
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-416

Query Match      3.8%; Score 57.5; DB 6; Length 708;
Best Local Similarity 21.9%; Pred. No. 1.8e+02;
Matches 57; Conservative 23; Mismatches 89; Indels 91; Gaps 12;

Qy 4 SEATPTSPGPFESKY-----PEFHGVRLLPPFCRGKWEI--ANFVRPSDVWIV 51
Db 268 SSVVTSLSGPAVKKHVGLLRIGKRMNMHKIPLHTVRQFFMEDIVLANHF-----DIFNP 323
Qy 52 TYPKSGTSLQEVVYLVVSGQADPEIGLNMNIDEQLPVLEYQPQGLDIIKELTSPRLIKSH 111
Db 324 DNPKV-TQAIQSFC-----LEKIEEMLENAE-----RERLGNSH 356
Qy 112 LPYR-----FLPSDLHNGDSKVIYMARNPDKLVVSYQFHR----- 147
Db 357 QPEKPLVRLRVVDYSGGFPFVLRFSQKFVDRVANPKDII--HPRHREQKEKTGEEINF 414
Qy 148 -SLRTMSYRGT-----FQFCRRFMNDKLYGSGWFEHVGQFWE-HRMDNSNVLF 194
Db 415 GKLIKPSGEGTLRVDELVKQYFQPAEKNVQLSLITERGMGEAVQGFVDVKEEKDAIBELV 474
Qy 195 KYEDMHRDLVTMVEQLARFL 214
Db 475 KYQ-----LEKTQRF 485

RESULT 55
US-10-504-120-21
; Sequence 21, Application US/10504120
; Publication No. US2006008829A1
; GENERAL INFORMATION:
; APPLICANT: Exelixis, Inc.
; TITLE OF INVENTION: MINRS AS MODIFIERS OF INSULIN RECEPTOR SIGNALING AND METHODS OF
; FILE OF INVENTION: USE
; FILE REFERENCE: EX03-003C-PC
; CURRENT APPLICATION NUMBER: US/10/504,120
; PRIOR FILING DATE: 2004-08-06
; PRIOR APPLICATION NUMBER: 60/354,824
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/358,217
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 60/358,189
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 60/358,126
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 60/358,995
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/358,756
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/358,765
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/359,531
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/360,222
; PRIOR FILING DATE: 2002-02-26
; PRIOR APPLICATION NUMBER: 60/360,224
; PRIOR FILING DATE: 2002-02-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 1763
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-10-504-120-21

Query Match      3.8%; Score 57.5; DB 6; Length 1763;
Best Local Similarity 20.1%; Pred. No. 7.4e+02;
Matches 54; Conservative 41; Mismatches 100; Indels 73; Gaps 13;

Qy 36 EEIANFVRPSDVW-----IVTYPKSGTSLQEVV-----YLVSQGADPD-EI 77
Db 158 EELADFVLQWMDVGLSSEFLVLVNLVKFNSCYLDEVIARMVQMICLLCVRTASSVDIEV 217
Qy 78 GLMNID-----BQLPVLEYQPQGLDIIKELTSP-----RLIKSHLPYRFLPSDL 121
Db 218 SLQVLDVAVCVNCLPASLPLFIVTLCTINVKELCPCKWKLMMNLGLTHLGHSAIYNMC 277
Qy 122 HNGDSKVIYMARNP--KDLV-----VSYQFHR--SLRTMSYRGTFQFCRRFM--NDKLG 171
Db 278 HLMDRA-YMEDAPLLRGAVFFVGMLWGAHRLYSLRN-SPTSVFSPFSYQAMACPNVVS 335
Qy 172 YGSWFEHVQEFWEHRMDSNVLF-----KYEDMHRDLVTMVEOLA 211
Db 336 YEIVLSITRLIKYRKELQVWAMDILLNIERLLOQLDLSPELTIIVHDLTTVEEL- 394
Qy 212 RFLGVSCDKAQLEALTEHCHQLVDCCN 239
Db 395 -----CDQNEFHGSGQERYFELVERCAD 416

RESULT 56
US-10-505-928-288
; Sequence 288, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 288
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-288

Query Match      3.7%; Score 57; DB 6; Length 720;
Best Local Similarity 20.3%; Pred. No. 2.2e+02;
Matches 56; Conservative 36; Mismatches 94; Indels 90; Gaps 14;

Qy 32 RGMKEETANFVRPSDVWIVTYPKSGTSLQEVVYLVVSGQADPEIGLNMNIDEQLPVL-- 89
Db 265 KASMENQPHSKPLPSDIAVIMYTSGSTCLPKGVMIHSNII-----AGITGMAERIPELGE 320
Qy 90 EYPQPG---LDIIKELTSPRIKSH---LPYRFLPSDLHNGDSKV----- 128
Db 321 EDVYIGYLPPLAHVLELSAELVCLSHGCRIGYS-SPQTLADQSSKIKGSKGDSKDTMLKPTL 379
Qy 129 -----IYMARNPKDLVVSYQFHRSLRTMSY-----RGTFQFCRRFMDNKL 170
Db 380 MAAVPEIMDRIV--KNVWNKYSEMSSFOQNLFILAYNYKMEQISKGRNTPLCDSFV---- 433
Qy 171 YGSWFEHVQEFWEHRMDSNVLFKYEDEMHRDLVTMVEQLARFLGVSCDKAQLEALTEHC 230
Db 434 -----FRKVRSL-----LGGNIRLL-----LCGGAPLSATTQRF 462
Qy 231 HQLVDQCCNAEALPVGRGRVGLWKDIFTVSNNEKFD 266
Db 463 MNIC-FCC-----PVQOG-YGLTESAGAGTISEVWD 491
```

```
RESULT 57
US-10-505-928-491
; Sequence 491, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: Patentin 3.2
; SEQ ID NO 491
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-491

Query Match      3.7%; Score 57; DB 6; Length 832;
Best Local Similarity 20.9%; Pred. No. 2.7e+02;
Matches 38; Conservative 29; Mismatches 57; Indels 58; Gaps 8;

QY 100 KELTSP-RLI-----KSHLPYRFLPSDLHGDSKVIYMARNPKDLVSVYYQFHRSLRTWS 153
Db 109 KEITLPSRLIYYINODSESPYHVL-----DTKARHQKHNAKAVHLAQASQF----- 154
QY 154 YRGTFQECRRFRFMDNDKLGYSWFHVFQEFWEHRMDSNVFLKYEDMHRDLVTMVQLARF 213
Db 155 -----IEAFGSKFILDLI-----LNNGLLSSDYVEIHYE----- 183
QY 214 LGVSCDQKALEALTEHC--HQLVDQCNAE-ALPVGGRGVGLWKDIFTVSMNEKFDLVYK 270
Db 184 -----NGKPYSGGEGHYHGSIRGVKDSKVALSTCNGLHGFEDDTFYVMIEPLELVHD 239
QY 271 QK 272
Db 240 EK 241

RESULT 58
US-10-505-928-35
; Sequence 35, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: Patentin 3.2
; SEQ ID NO 35
; LENGTH: 926
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-35

Query Match      3.7%; Score 57; DB 6; Length 926;
Best Local Similarity 28.4%; Pred. No. 3.2e+02;
Matches 21; Conservative 16; Mismatches 33; Indels 4; Gaps 3;

QY 5 EAETSTP--GFESKYFEFHGVRLLPFCRGKMBEIANFPVR-PSDVWLVTPKSGTSL 61
Db 249 KTEITLSKPDSKQSKKGLKFLRL-SFKKDKTKQLANFSQAQFPPEWPLRDEDTPATIP 307
QY 62 QEVVYLVSQADPD 75
Db 308 REVEWEIIRINPD 321

RESULT 59
US-10-505-928-100
; Sequence 100, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: Patentin 3.2
; SEQ ID NO 100
; LENGTH: 2871
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-100

Query Match      3.7%; Score 57; DB 6; Length 2871;
Best Local Similarity 20.4%; Pred. No. 1.7e+03;
Matches 47; Conservative 40; Mismatches 91; Indels 52; Gaps 11;

QY 1 MAESAEATPSTPGEPESKYFEFH--GVRLLPPFCRGK-----MEEIANFPVRPSDVWIVT 52
Db 2222 LVDSGILRPSTVNELESQISYDEVGERIKDFLQSSCIAGIYNETTQKGLGIYEAMKIG 2281
QY 53 YPKSGTSL-LQEVVYLVSQADPDPEIGLNMNIDEQLPVLEYPOPLDII-----KELTSPL 107
Db 2282 LVRPGTALELELAQAATGFIVDP-----VSNL--RLPVEEAYKRGVLGVIEFKEKLLSAERA 2335
QY 108 IKSHLPYRFLPSDLHGDSKVIYMARNPKDL-----VSVYYQFHRSL 149
Db 2336 VTGY-----NDPETHNIISLFOAMN-KELIEKHGIRLLEAQATGDIIDPKSHRLP 2387
QY 150 RTMSY-RGTQFQECRRFRFMDNDKLGYSWFHVFQEFWEHRMDSNVFLKYED 198
Db 2388 VDIAYKRGYFNEELSEILSDP-----SDTKGFFDPNTEENLTYLQKE 2431

RESULT 60
US-10-505-928-167
; Sequence 167, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: Patentin 3.2
; SEQ ID NO 167
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-167

Query Match      3.7%; Score 56.5; DB 6; Length 493;
Best Local Similarity 18.7%; Pred. No. 1.4e+02;
Matches 63; Conservative 43; Mismatches 108; Indels 123; Gaps 17;

QY 19 YFEFHGVRLLPFCRGKMBEIANFPVRPSDVWIVTPKSGTSLLEQVYLVSQADPDIG 78
Db 161 YLSFHKLL--HLQGERE-----PG--WI---KQLFTNFISFTLXLVLKGQICKEIN 205
QY 79 LM-NI-----DEQLPVLEYPOPLDIIKELTSPLI-----KSHLPYRFLPSDLH 122
Db 206 VISNIMADFVQTRAASILSDGIDIGVDI--SLTGDPFVITASYLESHHKGHFIYKNVSEDL 263
```

QY 123 -----NGDSKVIYMARPKDLVSVYQPHRSRLTMSYRGTFQFCRRFMNDKLGYS 174  
Db 264 LPTSPPTLLGSRMLY-----FWFSERVPHSLA-KVAFOD--GRMLSLUM--GD 307  
QY 175 WFEHQEFWEHRMDSNV-----LF 193  
Db 308 EFKAVLETWGFNTQEIHQEVVGGFPFQAQVTVHCLKMPKISQNKQGVVNVSSVMVKFLP 367  
QY 194 LKYEDMHRDLVTMVBQLARFLGVSCDKAQL-----BALTEHCHQLVD 235  
Db 368 PRPDQOHSVAYTFEEDIVTTVOASYSKKLFLSLDLDFQITPKTVSNLTSSESIQSFLQ 427  
QY 236 QCCNAEALPVGRGRVGLMKDIFTVMNEK----FDLV 268  
Db 428 SMITAVGPEVMSRLEV---VFTALMNSKGVSLFDII 461

## RESULT 61

US-10-511-937-2991  
; Sequence 2991, Application US/10511937  
; Publication No. US2006008836A1  
; GENERAL INFORMATION:  
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.  
; APPLICANT: Wohlgemuth, Jay  
; APPLICANT: Fry, Kirk  
; APPLICANT: Woodward, Robert  
; APPLICANT: Ly, Ngoc  
; APPLICANT: Prentice, James  
; APPLICANT: Morris, MacDonald  
; APPLICANT: Rosenberg, Steven  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING  
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION  
; FILE REFERENCE: 506612000104  
; CURRENT APPLICATION NUMBER: US/10/511,937  
; CURRENT FILING DATE: 2004-10-19  
; PRIOR APPLICATION NUMBER: PCT/US2003/012946  
; PRIOR FILING DATE: 2003-04-24  
; PRIOR APPLICATION NUMBER: US 10/131,831  
; PRIOR FILING DATE: 2003-04-24  
; PRIOR APPLICATION NUMBER: US 10/325,899  
; PRIOR FILING DATE: 2002-12-20  
; NUMBER OF SEQ ID NOS: 3117  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2991  
; LENGTH: 505  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-511-937-2991

Query Match 3.7%; Score 56.5; DB 6; Length 505;  
Best Local Similarity 29.3%; Pred. No. 1.4e+02;  
Matches 29; Conservative 9; Mismatches 22; Indels 39; Gaps 7;  
QY 37 EIANPVR-----PSDVW-----IVTY-----PKSGTSLQBVVYLVSGQ 71  
Db 394 EGAKFPKWTAPAINFGSFTIKSDVWSFGILLMEIVTYGRIPYFGMS-NPEVIRALERG 452  
QY 72 AD-----PDRI-GLM-----NIDEQLPVLEYQPGLD 97  
Db 453 YMRPRPENCPEELYNIMRCWKNRPEERTFEYIQSVLD 491

## RESULT 62

US-11-301-924-4  
; Sequence 4, Application US/11301924  
; Publication No. US20060090218A1  
; GENERAL INFORMATION:  
; APPLICANT: Joshua Z. Levin  
; APPLICANT: Ken Phillips  
; APPLICANT: Greg Budziszewski  
; APPLICANT: Fred Meins  
; APPLICANT: Zhenya Glazov

; TITLE OF INVENTION: Methods of Controlling Gene Expression  
; FILE REFERENCE: 31481USNP  
; CURRENT APPLICATION NUMBER: US/11/301,924  
; PRIOR FILING DATE: 2005-12-13  
; PRIOR APPLICATION NUMBER: US/09/896,186  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: 60/222,202  
; PRIOR FILING DATE: 2000-08-01  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 642  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-11-301-924-4

Query Match 3.7%; Score 56.5; DB 7; Length 642;  
Best Local Similarity 19.6%; Pred. No. 2.1e+02;  
Matches 45; Conservative 40; Mismatches 86; Indels 59; Gaps 11;  
QY 42 PVRSDWIVTYPKSGTSLQEVVY-LVSGQADPDEI--GLMNIDEQLPVLEYDQ----P 94  
Db 104 FVKPLP-----LEETPFKLVEEVKDLLEDAALQSVEEFAVDLEHNOYRTFQ 150  
QY 95 GLDIIKELTSPR-----LIKSHL-PY-RFLPSD-----LHNGDSKVIYMARNPXDL 138  
Db 151 GLTCLMQISTEDYIVDFKLWDHIGLYBELFKDKKKVIHGADEDIILWQDFGIY 210  
QY 139 VVSYYQPHRSRLTMSYRGTFQFCRRFMNDKLGYSMFHVQEFWE-----HRM 187  
Db 211 VCNLFDTQASRVLKLERNLSLEF---LLKHVCGVAANKKEYQKADWRIRPLPDVMKRYARE 267  
QY 188 DSNVFLKYEDMHRDLVTMVBQLARFLGVSCDKAQLAALTEHCHQLVDQC 237  
Db 268 DTHYLLYIYDVMRMELHTMAKE-----DEQSDSPLVEVYKRSYDVC 308

## RESULT 63

US-10-468-193-7  
; Sequence 7, Application US/10468193  
; Publication No. US20060100416A1  
; GENERAL INFORMATION:  
; APPLICANT: Palli, Subba R.  
; APPLICANT: Cress, Dean E.  
; APPLICANT: Fujimoto, Ted T.  
; APPLICANT: Kumar, Mohan B.  
; TITLE OF INVENTION: Novel Substitution Mutant Receptors and Their Use in A Nuclear  
; TITLE OF INVENTION: Receptor-Based Inducible Gene Expression System  
; FILE REFERENCE: A01247-US  
; CURRENT APPLICATION NUMBER: US/10/468,193  
; CURRENT FILING DATE: 2002-02-20  
; PRIOR APPLICATION NUMBER: US 60/313,925  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: PCT/US02/05090  
; PRIOR FILING DATE: 2002-02-20  
; NUMBER OF SEQ ID NOS: 113  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 7  
; LENGTH: 147  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-10-468-193-7

Query Match 3.7%; Score 56; DB 6; Length 147;  
Best Local Similarity 31.4%; Pred. No. 24;  
Matches 22; Conservative 16; Mismatches 30; Indels 2; Gaps 2;  
QY 84 EQLPVLEYQPGLDIIKELTSPRLIKSHLPYRFLPSPDLHNGDSKVIYMARNPXDLVSVY 143  
Db 65 EQLFLIIFREDLDMILKWDLSQDIKALLTGLFVQDNV-NKDAVTDRLASVETDPLTLR 123  
QY 144 QPHRSRLTMS 153

Db 124 Q-HRISATSS 132

RESULT 64

US-11-264-784-54

; Sequence 54, Application US/11264784

; Publication No. US20060094092A1

; GENERAL INFORMATION:

; APPLICANT: E.I. duPont de Nemours & Co., Inc.

; APPLICANT: Damude, Howard Glenn

; APPLICANT: Gillies, Peter John

; APPLICANT: Macool, Daniel Joseph

; APPLICANT: Picataggio, Stephen K.

; APPLICANT: Pollak, Dana M. Walters

; APPLICANT: Ragghianti, James John

; APPLICANT: Xue, Zhixiong

; APPLICANT: Yadav, Narendra S.

; APPLICANT: Zhang, Hongxiang

; APPLICANT: Zhu, Quinn

; TITLE OF INVENTION: HIGH ARACHIDONIC ACID PRODUCING STRAINS OF YARROWIA LIPOLYTICA

; FILE REFERENCE: CL3136 USA

; CURRENT APPLICATION NUMBER: US/11/264,784

; CURRENT FILING DATE: 2005-11-01

; NUMBER OF SEQ ID NOS: 375

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 54

; LENGTH: 275

; TYPE: PRT

; ORGANISM: Mortierella alpina

US-11-264-784-54

Query Match 3.7%; Score 56; DB 7; Length 275;

Best Local Similarity 21.7%; Pred. No. 64;

Matches 13; Conservative 16; Mismatches 27; Indels 4; Gaps 2;

QY 139 VVSYQFHRSLTMSYRGTFQFEC---RRFMNDKLGYSWFEHVOEFWEHRMDSNVLFK 195

Db 85 ITFYMFAMVAKNFRTHLHEAYCDTQSLNNAALGYWGLFYLSKPYE-VIDTIIILK 143

RESULT 65

US-10-489-730-4

; Sequence 4, Application US/10489730

; Publication No. US20060088825A1

; GENERAL INFORMATION:

; APPLICANT: Melino, Gennaro

; APPLICANT: Hayes, Ian

; APPLICANT: de Laurenzi, Vincenzo

; APPLICANT: Barcaroli, Daniela

; APPLICANT: Candi, Eleonora

; APPLICANT: Bernassola, Francesca

; APPLICANT: Tobler, Andreas

; APPLICANT: Novak, Urban

; TITLE OF INVENTION: Human Delta-N p73 Molecules and Uses Thereof

; CURRENT APPLICATION NUMBER: US/10/489,730

; CURRENT FILING DATE: 2004-03-16

; PRIOR APPLICATION NUMBER: PCT/GB02/04238

; PRIOR FILING DATE: 2002-09-17

; PRIOR APPLICATION NUMBER: US 60/322,436

; PRIOR FILING DATE: 2001-09-17

; NUMBER OF SEQ ID NOS: 39

; SEQ ID NO 4

; LENGTH: 450

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: peptide sequence of delta-N p73-beta protein

US-10-489-730-4

Query Match 3.7%; Score 56; DB 6; Length 450;

Best Local Similarity 29.7%; Pred. No. 1.4e+02;

Matches 22; Conservative 15; Mismatches 29; Indels 8; Gaps 4;

QY 62 QEVVYLVSGADPDDEIGLMNIDEQLPVL E-YPOQGLDIKELTSPLIKSHL-PYRFLP- 118

Db 303 EDTYYLVQVRGRENFEI-LMKLKELELMELVPQLVDSYRQQQLLQRPSPHLPQPSYGPV 361

QY 119 ---SDLHGDSKV 128

Db 362 LSPMKVHGGMNKL 375

RESULT 66

US-10-528-032-17

; Sequence 17, Application US/10528032

; Publication No. US20060101530A1

; GENERAL INFORMATION:

; APPLICANT: EXELIXIS, INC.

; TITLE OF INVENTION: RORS AS MODIFIERS OF THE p21 PATHWAY AND METHODS OF USE

; FILE REFERENCE: EX03-068C-US

; CURRENT APPLICATION NUMBER: US/10/528,032

; CURRENT FILING DATE: 2005-03-16

; PRIOR APPLICATION NUMBER: US 60/411,010

; PRIOR FILING DATE: 2002-09-16

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 17

; LENGTH: 459

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-528-032-17

Query Match 3.7%; Score 56; DB 6; Length 459;

Best Local Similarity 22.3%; Pred. No. 1.4e+02;

Matches 54; Conservative 37; Mismatches 85; Indels 66; Gaps 16;

QY 56 SGTSLQVWYLVSGADPDDEIGLMNIDEQLPVLEYPOQGLDI--IKEL-TSPRLIKSHL 112

Db 138 SGTANGHVIDL-----PKSEGYNVDSGQSPD--OSGLDMTGKIKQEPIDYDLTSV 189

QY 113 PYRFLPSDLHGD-----SKVIYMARPKDLVVSYQFHRSLTMSYRGTFQFECRRF 165

Db 190 PNLFTYSSFNNGQLAPGITMTETDRIAQN---IIKSH-----LTCQY--TMEE----- 233

QY 166 MNDKLGYSWFEHVOE-----FWHRMDSNVLFKYEDMHRDLVTWVQLARF 213

Db 234 ----LHQLAWQTHTYEIKAYQSKSREALWQCCAQIQTHTAIQY-----VVFEAKRITGF 283

QY 214 LGVSCDKAQLREALTEHQ--LVQOC-----CNAEALPVGR-GRVGLWKDIFTVSM-NEK 264

Db 284 MEL-CONDQILLKSGCLEVVLRMCRAFNPINNTLVTFEGKYGMQMPKALGSDDLVNEA 342

QY 265 FD 266

Db 343 FD 344

RESULT 67

US-11-301-554-336

; Sequence 336, Application US/11301554

; Publication No. US20060088527A1

; GENERAL INFORMATION:

; APPLICANT: Henderson, Robert A.

; APPLICANT: Wang, Tongtong

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Kalos, Michael D.

; APPLICANT: Sleath, Paul R.

; APPLICANT: Johnson, Jeffrey C.

; APPLICANT: Retter, Marc W.

; APPLICANT: Durham, Margarita

; APPLICANT: Carter, Darrick

; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: McNabb, Andria

```
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C21
; CURRENT APPLICATION NUMBER: US/11/301.554
; CURRENT FILING DATE: 2005-12-13
; PRIOR APPLICATION NUMBER: US 10/283.017
; PRIOR FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US 10/113.872
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 10/017.754
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 09/902.941
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 09/849.626
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: US 09/736.457
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 09/702.705
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: US 09/677.419
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 09/671.325
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 09/658.824
; PRIOR FILING DATE: 2000-09-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2157
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 336
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-301-554-336

Query Match          3.7%; Score 56; DB 7; Length 480;
Best Local Similarity 22.1%; Pred. No. 1.5e+02;
Matches 48; Conservative 29; Mismatches 80; Indels 60; Gaps 12;

Qy 29 PFCR---GKMEIANFPVPDVMIVTPYKSGTSLQSVVLSVSGADPDIGLNMIDEQ 85
Db 86 PGCSSLDGLLTHGHPFLVQPGVTLVYNPYS-WNLIANVLYLES-----PAGVGFSYDDK 140

Qy 86 LPV---LEVPOGLDIIKELTSPRLKSHLPYRPLPSDLHN-----GDSKV-IYM----- 131
Db 141 FYATNDTEVAGSNFALQDF-----FRLEPEYKNNKLFITGSEYAGIYIPTLAV 189

Qy 132 -----ARNPKDLVV-----SYQPHRSLRTMSYRGTFQFCRRPMNDKLGYSWFHVQ 180
Db 190 LVMQDPSMNLQGLAVGNGLSSYEQNDNSLVVFAY-----YHGLLGNRLW-----S 234

Qy 181 EFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVS 217
Db 235 SLQTHCCSQNKCNF-YDNKDLCEVTNLQEVARIQVNS 270

RESULT 68
US-09-949-925-82
; Sequence 82, Application US/09949925
; Publication No. US20060099575A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 67 Human secreted proteins
; FILE REFERENCE: P2032P2
; CURRENT APPLICATION NUMBER: US/09/949,925
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US 60/232,150
; PRIOR FILING DATE: 2000-12-09
; PRIOR APPLICATION NUMBER: PCT/US99/01621
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: US 60/073,160
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: US 60/073,159
; PRIOR FILING DATE: 1998-01-30

; PRIOR APPLICATION NUMBER: US 60/073,165
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: US 60/073,164
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: US 60/073,167
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: US 60/073,162
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: US 60/073,161
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: US 60/073,170
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 82
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (326)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (342)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (444)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-949-925-82

Query Match          3.7%; Score 56; DB 1; Length 578;
Best Local Similarity 18.4%; Pred. No. 2e+02;
Matches 35; Conservative 35; Mismatches 74; Indels 46; Gaps 9;

Qy 35 MEEIANFPVPDVMIVTPYKSGTSLQSVVLSVSGADPD--BIGLMNIDEQLPVLEYP 92
Db 361 VEKREGYPLRP-----ELIESAMVLYRATGCDPTLLELG-RDAVESIEKISKV 406

Qy 93 QPGLDIIKELTSPRLKSHLPYRPLPSDLHNGDSKVLYMARNPKDLVVSYVQPHRSRLTM 152
Db 407 ECGPATTKDURDHKL-DNRMESFFLAETV-----KYLLYLLFDPXNFI-----HNN----- 450

Qy 153 SYRGTFQFCRRFMNDKLGYSWFHVQEFWEHRMDSNVLF-----AHPIDPALHCCORLKEQWEVEDLMREF 203
Db 451 --GSTFDVAVTPYCECILGAGGIYFNTE---KYLILFDPXNFI-----HNN----- 450

Qy 204 VTWVEQLARF 213
Db 506 YSLKRSRSKF 515

RESULT 69
US-10-489-730-2
; Sequence 2, Application US/10489730
; Publication No. US20060088825A1
; GENERAL INFORMATION:
; APPLICANT: Mellino, Gennaro
; APPLICANT: Hayes, Ian
; APPLICANT: de Laurenzi, Vincenzo
; APPLICANT: Barcaroli, Daniela
; APPLICANT: Candi, Eleonora
; APPLICANT: Bernassola, Francesca
; APPLICANT: Tobler, Andreas
; APPLICANT: Novak, Urban
; TITLE OF INVENTION: Human Delta-N p73 Molecules and Uses Thereof
; FILE REFERENCE: 19319.002
; CURRENT APPLICATION NUMBER: US/10/489,730
; CURRENT FILING DATE: 2004-03-16
; PRIOR APPLICATION NUMBER: PCT/GB02/04238
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: US 60/322,436
; PRIOR FILING DATE: 2001-09-17
; NUMBER OF SEQ ID NOS: 39
```



```
; SEQ ID NO 2
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: peptide sequence of delta-N p73-alpha protein
US-10-489-730-2

Query Match          3.7%; Score 56; DB 6; Length 587;
Best Local Similarity 29.7%; Pred. No. 2e+02;
Matches 22; Conservative 15; Mismatches 29; Indels 8; Gaps 4;

Qy 62 QEVYLVSQGADPDIGLWNIDEQLPVLE-YPOQGLDIKELTSRLIKSHL-PYRFLP- 118
:: || || || || || || || || || || || || || || || || || || ||
Db 303 ETTYLVQVRGRENFEI-LMKLKESELMLVLPQPLVDYSYRQQQLQRPShLQPPSYG 361
:: || || || || || || || || || || || || || || || || || || ||

Qy 119 ----SDLHNGDSKV 128
: : | | | :
Db 362 LSPMKVHVGGMNKL 375
: : | | | :

RESULT 70
US-10-489-730-11
; Sequence 11, Application US/10489730
; Publication No. US20060088825A1
; GENERAL INFORMATION:
; APPLICANT: Melino, Gennaro
; APPLICANT: Hayes, Ian
; APPLICANT: de Laurenzi, Vincenzo
; APPLICANT: Barcaroli, Daniela
; APPLICANT: Candi, Eleonora
; APPLICANT: Bernassola, Francesca
; APPLICANT: Tobler, Andreas
; APPLICANT: Novak, Urban
; TITLE OF INVENTION: Human Delta-N p73 Molecules and Uses Thereof
; CURRENT APPLICATION NUMBER: US/10/489,730
; CURRENT FILING DATE: 2004-03-16
; PRIOR APPLICATION NUMBER: PCT/GB02/04238
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: US 60/322,436
; PRIOR FILING DATE: 2001-09-17
; NUMBER OF SEQ ID NOS: 39
; SEQ ID NO 11
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Reported sequence of TA p73 from GenBank Accession AL136528
US-10-489-730-11

Query Match          3.7%; Score 56; DB 6; Length 661;
Best Local Similarity 29.7%; Pred. No. 2.4e+02;
Matches 22; Conservative 15; Mismatches 29; Indels 8; Gaps 4;

Qy 62 QEVYLVSQGADPDIGLWNIDEQLPVLE-YPOQGLDIKELTSRLIKSHL-PYRFLP- 118
:: || || || || || || || || || || || || || || || || || || ||
Db 377 ETTYLVQVRGRENFEI-LMKLKESELMLVLPQPLVDYSYRQQQLQRPShLQPPSYG 435
:: || || || || || || || || || || || || || || || || || || ||

Qy 119 ----SDLHNGDSKV 128
: : | | | :
Db 436 LSPMKVHVGGMNKL 449
: : | | | :

RESULT 71
US-10-511-455-56
; Sequence 56, Application US/10511455
; Publication No. US20060088835A1
; GENERAL INFORMATION:
; APPLICANT: Pickard, Benjamin Simon
; APPLICANT: Blackwood, Douglas
; APPLICANT: Porteous, David
; APPLICANT: Muir, Walter John
```

```
; APPLICANT: Mors, Ole
; APPLICANT: Ewald, Henrik Lykke
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES
; FILE REFERENCE: 9013.63
; CURRENT APPLICATION NUMBER: US/10/511,455
; CURRENT FILING DATE: 2004-10-05
; PRIOR APPLICATION NUMBER: PCT/GB03/001543
; PRIOR FILING DATE: 2003-04-07
; PRIOR APPLICATION NUMBER: GB0207902.8
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0207904.4
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0207900.2
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0207901.0
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: GB0227734.1
; PRIOR FILING DATE: 2002-11-28
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 56
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-455-56

Query Match          3.7%; Score 56; DB 6; Length 736;
Best Local Similarity 25.2%; Pred. No. 2.9e+02;
Matches 39; Conservative 20; Mismatches 54; Indels 42; Gaps 10;

Qy 25 VRLPPFCRGKMEETANFP-----VRPSDVWIVTPKSGTSLLOEVVYLVSQGADPD 79
: || | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 50 LQLPPLSQ-RQSERARTPEGDGSRPTTLPILTLPSTIAITTVSQECFVENGPSP---GR 105
: || | : | | : | | : | | : | | : | | : | | : | | : | | : | | :

Qy 80 MNIDEQLP-----VLEYPQGLDIKEL-----TSPRLIKSHLPYRFLPSDLHNG 124
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 106 SPLDPQASSSAGLVHATFPGHRSQRRESFLYRSDSDYDLSPKAMSRN---SSLPSEQH-G 161
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :

Qy 125 DSKVIYMARNPKDLVVS-YYQFHRSLRTMSYRGTF 158
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 162 D-----DLIVTPFAQVLASLR---SVRNPF 183
: | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :

RESULT 72
US-11-219-635-3
; Sequence 3, Application US/11219635
; Publication No. US20060100218A1
; GENERAL INFORMATION:
; APPLICANT: IBRAHIM, PRABHA
; APPLICANT: BREMER, RYAN
; APPLICANT: GILLETTE, SAM
; APPLICANT: CHO, HANNA
; APPLICANT: NESPI, MARIKA
; APPLICANT: MANO, SHUMEYE
; APPLICANT: ZHANG, CHAO
; APPLICANT: ARTIS, DEAN R.
; APPLICANT: LEE, BYUNGHUN
; APPLICANT: ZUCKERMAN, REBECCA
; TITLE OF INVENTION: PDE4B INHIBITORS
; FILE REFERENCE: 039363-1111
; CURRENT APPLICATION NUMBER: US/11/219,635
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: 60/607,407
; PRIOR FILING DATE: 2004-09-03
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent in ver. 3.3
; SEQ ID NO 3
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-219-635-3

Query Match          3.7%; Score 56; DB 7; Length 736;
```







```
; APPLICANT: Durham, Margarita
; APPLICANT: Carter, Darick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C21
; CURRENT APPLICATION NUMBER: US/11/301,554
; CURRENT FILING DATE: 2005-12-13
; PRIOR APPLICATION NUMBER: US 10/283,017
; PRIOR FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US 10/113,872
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 10/017,754
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 09/902,941
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 09/849,626
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: US 09/736,457
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 09/702,705
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: US 09/677,419
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 09/671,325
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 09/658,824
; PRIOR FILING DATE: 2000-09-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2157
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1903
; LENGTH: 2209
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-301-554-1903

Query Match      3.6%; Score 55.5; DB 7; Length 2209;
Best Local Similarity 20.7%; Pred. No. 1.7e+03;
Matches 46; Conservative 40; Mismatches 99; Indels 37; Gaps 11;

Qy 24 GVRLLPFCRGKMEETANFVRPSDVWIVTY-PKSGTSLLOEVVYLVSGQADPDEIGLMI 82
Db 1988 GFNMIPYLRKVLKAISSI-----HSLWQVPYFSKAWQRVQIPLLSASCPLSPDQ--LSDC 2041

Qy 83 DEQL-PVLEYQPG-LDIKELTSPRLIKSHLP-----YRFLP-SDLHNGDSKVIYMAR 133
Db 2042 SESLAVLECPVSGDLDLIG--VARQYIQLELPAPAFALACLMLMPHSEKRHOQIKNFLGSC 2099

Qy 134 NPKDLVSVYQFHRSLRTMSYRGTFQFCRRFMND-----KLGYGSWFHEHVQEFWEHRMD 188
Db 2100 DPQ-----VLKQLEEHMTGQLAGFSHQIRSLILNINIKKFGIILAKTKYQMLKQHANV 2156

Qy 189 SNVLFLKYEDKMRDLVMTVEQQLARELVGSCDKAQLEALTEHC 230
Db 2157 TN-----NITELVNVYLANDLSLDEASVLITEYSKHC 2187

RESULT 84
US-10-511-937-2546
; Sequence 2546, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, Macdonald
```

```
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2546
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2546

Query Match      3.6%; Score 55; DB 6; Length 274;
Best Local Similarity 22.5%; Pred. No. 82;
Matches 40; Conservative 19; Mismatches 59; Indels 60; Gaps 9;

Qy 1 MAESEAEPTSTPGFEESKYFEFHGVRLE-----PFCRGKMEETANFVRPSDVWIVTYP 54
Db 105 LAQLDGEDPCT--EFRQQQLTVHRVHTFLPHEPP-----PPRPHDVTIVA-- 148

Qy 55 KSGTSLLOEVVYLVSGQADPDEIGLMI 94
Db 149 QLSWDRLOMLCALCRHWGPMNSLALYLTDAEAQQLHFVEASPVLAARQDVAYHVYVREG 208

Qy 95 GLDIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPCKDLVSVY--YQFHRSLR 150
Db 209 PLYPNQLRNVALAQALTPYVFL-SDI-----DPLPAYSLYDYLREAR 250

RESULT 85
US-11-283-522-35
; Sequence 35, Application US/11283522
; Publication No. US20060088914A1
; GENERAL INFORMATION:
; APPLICANT: Adams, Sean
; APPLICANT: Pan, James
; TITLE OF INVENTION: UCP5
; FILE REFERENCE: P1663R2
; CURRENT APPLICATION NUMBER: US/11/283,522
; CURRENT FILING DATE: 2005-11-18
; PRIOR APPLICATION NUMBER: US/10/270,861
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US/09/433,622
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: US 60/110,286
; PRIOR FILING DATE: 1998-11-30
; PRIOR APPLICATION NUMBER: US 60/129,583
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: US 60/143,886
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 35
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-283-522-35

Query Match      3.6%; Score 55; DB 7; Length 300;
Best Local Similarity 25.2%; Pred. No. 94;
Matches 30; Conservative 17; Mismatches 44; Indels 28; Gaps 6;

Qy 97 DIIKE-LTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPCKDLVSVY-----QFHRSL 149
Db 189 DILKEKLLDYHLLTDNFPCHFVSA---FGAGFCATVVASPVDVVKTRMNSPPQGYFSPL 245
```

Qy 150 RTMSYRGTFQF-----FCRFRMNDKLGYSWFHVFQFWEHRMDSNVLFKLYEDMHRDLV 204  
Db 246 DCM-1KMWAGQPTAFYKGFPTSPFLRGLSW-----NVVMFVTYEQKRALM 290

RESULT 86

US-10-511-937-2511  
; Sequence 2511, Application US/10511937  
; Publication No. US2006008836A1  
; GENERAL INFORMATION:  
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.  
; APPLICANT: Wohlgenuth, Jay  
; APPLICANT: Fry, Kirk  
; APPLICANT: Woodward, Robert  
; APPLICANT: Ly, Ngoc  
; APPLICANT: Prentice, James  
; APPLICANT: Morris, MacDonald  
; APPLICANT: Rosenberg, Steven  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING  
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION  
; FILE REFERENCE: 506612000104  
; CURRENT APPLICATION NUMBER: US/10/511.937  
; CURRENT FILING DATE: 2004-10-19  
; PRIOR APPLICATION NUMBER: PCN/US2003/012946  
; PRIOR FILING DATE: 2003-04-24  
; PRIOR APPLICATION NUMBER: US 10/131,831  
; PRIOR FILING DATE: 2002-04-24  
; PRIOR APPLICATION NUMBER: US 10/325,899  
; PRIOR FILING DATE: 2002-12-20  
; NUMBER OF SEQ ID NOS: 3117  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2511  
; LENGTH: 343  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-511-937-2511

Query Match 3.6%; Score 55; DB 6; Length 343;  
Best Local Similarity 27.2%; Pred. No. 1.2e+02;  
Matches 25; Conservative 11; Mismatches 28; Indels 28; Gaps 5;

Qy 170 LGYSWFE-----HVFQFWEHRM-DSNVLFKY-----EDMHRDLVTMVEQL 210  
Db 19 LNEGAWSVGALMTSPALLGYLQVCRARLPDSLMRKHVSVRQEDLQVRLSRPEAV 78  
Qy 211 ARFLGVSCDKAQLEALTEHCHQVLDQCCNAFA 242  
Db 79 AE---VKSFLIQLEAF-----LSRLCCTCEA 101

RESULT 87

US-11-305-413-13  
; Sequence 13, Application US/11305413  
; Publication No. US20060099699A1  
; GENERAL INFORMATION:  
; APPLICANT: Clausen, Henrik  
; APPLICANT: Amado, Margarita  
; TITLE OF INVENTION: UDP-GALACTOSE: BETA-N-ACETYL-GLUCOSAMINE BETA 1,3  
; TITLE OF INVENTION: GALACTOSYLTRANSFERASES, BETAGAL-T5  
; FILE REFERENCE: 7188-157  
; CURRENT APPLICATION NUMBER: US/11/305,413  
; CURRENT FILING DATE: 2005-12-16  
; PRIOR APPLICATION NUMBER: US/10/777,828  
; PRIOR FILING DATE: 2004-02-12  
; PRIOR APPLICATION NUMBER: US/09/831,630  
; PRIOR FILING DATE: 2001-05-10  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 378  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-305-413-13

Query Match 3.6%; Score 55; DB 7; Length 378;  
Best Local Similarity 19.7%; Pred. No. 1.3e+02;  
Matches 36; Conservative 15; Mismatches 44; Indels 88; Gaps 8;  
Qy 121 LHNGDSKVIYMAR-----NPKDLVSVYYQFHRSLRTMSYRGTFQFCCRRFMNDKLGYSW 175  
Db 217 LHSEVPLLYLGRVHWRVNS-----RTPGGRGVSE----- 248  
Qy 176 FEHVQEFWEH-----RMSNVLFKLYEDMHRDLVTMVEQL 210  
Db 249 -----EQPHTWGPFPYASGTGYLSASAVQLILKVASRAPLPLEDV----- 292  
Qy 211 ARFLGVSCDK-----AQLEALTEHCHQVLDQCCNAEALPVGRGRVGLWKDIFTVSMNEKF 265  
Db 293 --FVGSARRGGLAPTOCVKLAGATHYPLDRCCYGFLLTSH-RDDPWK-----MQEAW 343  
Qy 266 DLV 268  
Db 344 KLV 346

RESULT 88

US-10-196-749-152  
; Sequence 152, Application US/10196749  
; Publication No. US20060094864A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3430R1C340  
; CURRENT APPLICATION NUMBER: US/10/196,749  
; CURRENT FILING DATE: 2002-07-16  
; PRIOR APPLICATION NUMBER: 10/052586  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059266  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063120  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063121  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063486  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/063540  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063541  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063544  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 152  
; LENGTH: 541  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-196-749-152

Query Match 3.6%; Score 55; DB 6; Length 541;  
Best Local Similarity 18.4%; Pred. No. 2.3e+02;







```

Db      241 EKESAS-----ELHKLQFLLRRKAEVAT--ELPKTEVVIY--HGMSALQ 286
QY      63 EYVLVSOGADPD-----EIGLMNIDEQL-----EYPOPLDIIKELTSP 105
Db      287 KKYKAILMKDLADAPENETAKVKQLNLSQLRKCVDPYLPDGVPEPEP-FEVDHLTEA 345
QY      106 RLIIKSHLPYRFLPSDLHNGDSKVIYARNPK--DLVVSYYQFHRSLRTWTSYRG 156
Db      346 S-GKLHLLDKLL-AFLYSGHRVLLFSQMTOMLIDLQDY-----MDYRG 387

RESULT 95
US-10-505-928-359
; Sequence 359, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; PRIOR FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: Patentin 3.2
; SEQ ID NO 359
; LENGTH: 1015
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-359

Query Match      3.6%; Score 54.5; DB 6; Length 1015;
Best Local Similarity 19.5%; Pred. No. 6.8e+02;
Matches 43; Conservative 33; Mismatches 70; Indels 75; Gaps 10;

QY      3 ESEATPTSPGFEFESKYFEFHGVRLPPFCRGKMEIANFPVRPSDV-----WIVT-----Y 53
Db      472 ELQNMQPGSKEQS-----LPAGAQAELSDGLQLEVQFSEEAARGYIVTRDPLR 522
QY      54 PKSGTSLLOEVVYLV-----SQGADPEIG-----LMNIDQLPVLEYPQPGLDII 99
Db      523 PEGRRLEDVARLLQVSSAFADVEVLGPATVFKVSANQVNTTE-DVEKATVNDKXL 581
QY      100 KELTSPRLIK-----SHLPYRFLPSDLHNGDSK-----TFQEFRRFM 134
Db      582 BETSGLKULQTVGSKSLKFLPPQAEQEDSTKFIALTVLSLACILGVLASGLIVCLR 641
QY      135 PKDLVVSYYQFHRSLRTWTSYRG-----TFQEFRRFM 166
Db      642 SSQ-----HRLKELSLGLGGDFGADATAYQELCRQM 674
```

```

RESULT 96
US-11-311-778-24
; Sequence 24, Application US/11311778
; Publication No. US20060095988A1
; GENERAL INFORMATION:
; APPLICANT: Fischhoff, David A.
; TITLE OF INVENTION: SYNTHETIC PLANT GENES AND METHOD FOR PREPARATION
; FILE REFERENCE: 38-21 (13553) C
; CURRENT APPLICATION NUMBER: US/11/311,778
; CURRENT FILING DATE: 2005-12-19
; PRIOR APPLICATION NUMBER: 08/434,105
; PRIOR FILING DATE: 1995-05-03
; PRIOR APPLICATION NUMBER: 07/959,506
; PRIOR FILING DATE: 1992-10-09
; PRIOR APPLICATION NUMBER: 07/476,661
; PRIOR FILING DATE: 1990-02-12
; PRIOR APPLICATION NUMBER: 07/315,355
; PRIOR FILING DATE: 1989-02-24
; NUMBER OF SEQ ID NOS: 43
```

```

; SOFTWARE: Patentin version 3.1
; SEQ ID NO 24
; LENGTH: 1177
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: synthetic structural gene encoding a fusion protein derived from
; OTHER INFORMATION: B.t.k. HD-1 and B.t.k. HD-73
US-11-311-778-24
```

```

Query Match      3.6%; Score 54.5; DB 7; Length 1177;
Best Local Similarity 28.8%; Pred. No. 8.5e+02;
Matches 23; Conservative 8; Mismatches 32; Indels 17; Gaps 4;

QY      7 ETPSTPGFEFESKYFE---FHGVRLPP---FCRGKME-----EIANFPVRPSDVI 50
Db      1098 EAPSPADYASVYEKSYTDGRRENPCFENRGYRDYTLPLVGYYVTKLEYFP-ETDKWI 1156
QY      51 VTYPKSGTSLLOEVVYLVSQ 70
Db      1157 EIGETEGTFIVDSVELLME 1176
```

```

RESULT 97
US-11-311-778-10
; Sequence 10, Application US/11311778
; Publication No. US20060095988A1
; GENERAL INFORMATION:
; APPLICANT: Perlak, Frederick J.
; TITLE OF INVENTION: SYNTHETIC PLANT GENES AND METHOD FOR PREPARATION
; FILE REFERENCE: 38-21 (13553) C
; CURRENT APPLICATION NUMBER: US/11/311,778
; CURRENT FILING DATE: 2005-12-19
; PRIOR APPLICATION NUMBER: 08/434,105
; PRIOR FILING DATE: 1995-05-03
; PRIOR APPLICATION NUMBER: 07/959,506
; PRIOR FILING DATE: 1992-10-09
; PRIOR APPLICATION NUMBER: 07/476,661
; PRIOR FILING DATE: 1990-02-12
; PRIOR APPLICATION NUMBER: 07/315,355
; PRIOR FILING DATE: 1989-02-24
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 10
; LENGTH: 1178
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: synthetic structural gene encoding an insecticidal protein derive
; OTHER INFORMATION: d from B.t.k. HD-73
US-11-311-778-10
```

```

Query Match      3.6%; Score 54.5; DB 7; Length 1178;
Best Local Similarity 28.8%; Pred. No. 8.5e+02;
Matches 23; Conservative 8; Mismatches 32; Indels 17; Gaps 4;

QY      7 ETPSTPGFEFESKYFE---FHGVRLPP---FCRGKME-----EIANFPVRPSDVI 50
Db      1099 EAPSPADYASVYEKSYTDGRRENPCFENRGYRDYTLPLVGYYVTKLEYFP-ETDKWI 1157
QY      51 VTYPKSGTSLLOEVVYLVSQ 70
Db      1158 EIGETEGTFIVDSVELLME 1177

RESULT 98
US-11-311-778-12
; Sequence 12, Application US/11311778
; Publication No. US20060095988A1
; GENERAL INFORMATION:
; APPLICANT: Perlak, Frederick J.
```



GenCore version 5.1.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2006, 10:59:35 ; Search time 40 Seconds  
(without alignments)  
683.139 Million cell updates/sec

Title: US-10-768-158-2  
Perfect score: 1530  
Sequence: 1 MAESEAEPTSPGSESKYF.....FDLVYKQKMGKDLTFDFYL 284

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : PIR 80:\*

1: Pirl:\*  
2: Pirl:\*  
3: Pirl:\*  
4: Pirl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	495	32.4	298	2 S28183	aryl sulfotransfer
2	472	30.8	295	2 A55451	aryl sulfotransfer
3	467	30.5	296	2 JC7282	hydroxyarylamine s
4	458	29.9	291	2 S10329	aryl sulfotransfer
5	456	29.8	295	2 S52791	aryl sulfotransfer
6	456	29.8	295	2 JC5249	aryl sulfotransfer
7	456	29.8	295	2 G01843	aryl sulfotransfer
8	454	29.7	295	2 I57945	phenol-sulfating p
9	451	29.5	295	2 S52794	aryl sulfotransfer
10	451	29.5	295	2 JC5248	aryl sulfotransfer
11	449	29.3	295	2 JC5253	aryl sulfotransfer
12	447	29.2	296	2 JC7283	hydroxyarylamine s
13	446.5	29.2	299	2 JC5884	thyroid hormone su
14	446	29.2	295	2 G02924	phenol sulfotransf
15	446	29.2	295	2 S29045	estrone sulfotrans
16	444	29.0	295	2 S52399	aryl sulfotransfer
17	440.5	28.8	296	2 JC5885	thyroid hormone su
18	435	28.4	295	2 I73679	estrogen sulfotran
19	434	28.4	294	2 JC2229	estrogen sulfotran
20	432.5	28.3	299	2 JE0186	sulfotransferase (
21	430	28.1	295	2 I53296	testis-specific es
22	430	28.1	295	2 I56606	estrogen sulfotran
23	429	28.0	295	2 A41930	estrone sulfotrans
24	419.5	27.4	304	2 A49098	N-hydroxyarylamine
25	413	27.0	295	2 JE0197	phenol sulfotransf
26	412	26.9	294	2 JC5000	aryl sulfotransfer
27	408.5	26.7	296	2 A44011	adrenocortical est
28	393.5	25.7	301	2 JM0078	amine sulfotransf
29	385.5	25.2	287	2 JC4531	alcohol sulfotrans

30	380	24.8	338	2	JE0196	hydroxysteroid sul
31	376.5	24.6	287	2	A54026	alcohol sulfotrans
32	371	24.2	286	2	JE0152	alcohol sulfotrans
33	361.5	23.6	285	2	I65760	alcohol sulfotrans
34	359.5	23.5	284	2	I60190	alcohol sulfotrans
35	356.5	23.3	285	2	T10086	alcohol sulfotrans
36	355.5	23.2	285	1	I13848	alcohol sulfotrans
37	353.5	23.1	284	2	I52848	alcohol sulfotrans
38	351.5	23.0	284	2	A33569	alcohol sulfotrans
39	338	22.1	282	2	A26136	senescence marker
40	303	19.8	311	2	B40216	flavonol 3'-sulfot
41	302	19.7	329	2	T47448	sulfotransferase-1
42	297.5	19.4	324	2	B84452	probable steroid s
43	284.5	18.6	323	2	T47447	sulfotransferase-1
44	280	18.3	325	2	T07833	probable steroid s
45	278.5	18.2	346	2	E86319	probable flavonol
46	277.5	18.1	326	2	A84452	probable steroid s
47	276.5	18.1	323	2	T07831	probable steroid s
48	275.5	18.0	320	2	A40216	flavonol 4'-sulfot
49	267	17.5	350	2	H96768	protein flavonol s
50	265.5	17.4	324	2	T07832	probable steroid s
51	262.5	17.2	326	2	F86407	probable sulfotran
52	259.5	17.0	302	2	S69188	probable flavonol
53	254.5	16.6	296	2	JC7921	cytosolic sulfotra
54	249.5	16.3	338	2	A96769	protein flavonol s
55	240	15.7	333	2	A84523	probable steroid s
56	235.5	15.4	314	2	T06012	hypothetical prote
57	224.5	14.7	331	2	H84451	probable steroid s
58	161	10.5	280	2	A95936	probable alcohol s
59	110	7.2	249	2	E84072	hypothetical prote
60	109.5	7.2	261	2	G70393	conserved hypotet
61	106.5	7.0	326	2	B70958	hypothetical prote
62	101	6.6	5107	2	T29144	partial CDS - Caen
63	100.5	6.6	1424	2	T03851	thyroid hormone re
64	96	6.3	481	2	T27665	hypothetical prote
65	94.5	6.2	678	2	B64592	hypothetical prote
66	93	6.1	363	2	S18765	Sd protein - fruit
67	92.5	6.0	500	2	T03387	polyamine oxidase
68	92	6.0	759	2	T39468	anthranilate synth
69	91.5	6.0	1434	2	G71232	hypothetical prote
70	91	5.9	285	2	S49879	hypothetical narbo
71	91	5.9	285	2	S49898	hypothetical narbo
72	91	5.9	508	2	A71453	hypothetical prote
73	91	5.9	759	2	T46566	anthranilate synth
74	90.5	5.9	882	2	A57169	(heparan sulfate)-
75	90	5.9	435	2	D72379	hypothetical prote
76	89	5.8	332	2	B99791	hypothetical prote
77	88.5	5.8	882	2	A42855	N-heparan sulfate
78	88	5.8	604	2	C57784	MTG8 protein aplic
79	88	5.8	1752	4	A57784	AMU1/MTG8 mutant f
80	88	5.8	1767	2	T00458	hypothetical prote
81	87.5	5.7	254	2	B84071	hypothetical prote
82	87.5	5.7	308	2	C95935	probable sulfotran
83	87.5	5.7	466	2	H98017	conserved hypotet
84	87.5	5.7	883	2	A49733	(heparan sulfate)-
85	87.5	5.7	976	2	C84500	probable retroelem
86	87.5	5.7	1455	2	E75199	DNA-directed DNA p
87	86.5	5.7	434	2	A95150	conserved hypotet
88	86	5.6	836	1	JDVLD	DNA-directed DNA p
89	86	5.6	1452	2	T29107	polymorphic antige
90	85.5	5.6	352	2	A9298	conserved hypotet
91	85.5	5.6	1462	2	T42639	glucocorticoid rec
92	85	5.6	404	2	A34919	type II site-speci
93	85	5.6	463	2	S75889	hypothetical prote
94	85	5.6	607	1	J01535	hemagglutinin - ph
95	85	5.6	1105	2	S76557	carbamoyl-phosphat
96	84.5	5.5	782	2	D84514	dynamun-like prote
97	84	5.5	2747	2	B49132	fat facets (faf) s
98	83.5	5.5	419	2	C81744	ATP-dependent Clp
99	83.5	5.5	778	2	A86928	probable ATP-depen
100	83.5	5.5	1071	2	T22327	hypothetical prote



**Qy** 243 LP-VGRGVLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280  
| : | | | | | | | | | | | | | | | | | |  
**Db** 254 -PFMRKGWAGDWKTFTTVAQNEREDADYAERKAGCSLSF 291  
| : | | | | | | | | | | | | | | | | | |

**RESULT 3**

**JC7282**  
hydroxylamine sulfotransferase (EC 2.8.2.-) 2 - rat  
N;Alternate names: SULT1C2 sulfotransferase  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: JC7282  
R;Li, X.; Joehnk, C.; Hartmann, D.; Schestag, F.; Kroemer, W.; Gieselmann, V.  
Biochem. Biophys. Res. Commun. 272, 242-250, 2000  
A;Title: Enzymatic properties, tissue-specific expression, and lysosomal location of two  
A;Reference number: JC7282  
A;Contents: Kidney  
A;Accession: JC7282  
A;Molecule type: mRNA  
A;Residues: 1-296 <LIA>  
A;Cross-references: UNIPROT:Q9UWU8; UNIPARC:UIP00001360FC; GB:AJ238391  
C;Genetics:  
A;Gene: sult1c2  
C;Superfamily: alcohol sulfotransferase  
C;Keywords: kidney; sulfotransferase

Query Match 30.5%; Score 467; DB 2; Length 296;  
Best Local Similarity 35.1%; Pred. No. 9.7e-33;  
Matches 100; Conservative 58; Mismatches 105; Indels 22; Gaps 5;

**Qy** 16 ESKYFEFGVRLPPFCRGKEEIANFPVRPSDVMIIVTPKSGTSLLOEVVYLVSQGADPD 75  
| : | | | | | | | | | | | | | | | | | |  
**Db** 10 QTKLEKAVIGPLQAPTVDNNWSQIQTFRAKPDDLICTYPKSGTTWIEIVDMIEQNGDVE 69  
| : | | | | | | | | | | | | | | | | | |

**Qy** 76 EIGLMNTDEQLPVLEY---POP-GLDIIKELTSPLRIKSHLPYRFIPSDHLHGDSKVIM 131  
| : | | | | | | | | | | | | | | | | | |  
**Db** 70 KCQRTIIOHRHFFIEWARPQPQSGVDKANAMPAPRIILRTLPTQLLPSPFWTNCKFLYV 129  
| : | | | | | | | | | | | | | | | | | |

**Qy** 132 ARNPKDLDVSYQYQHSRLRTWSYRGTFQEFRCRRPMNDKLGVGWFHFQVFEHRMDNSV 191  
| : | | | | | | | | | | | | | | | | | |  
**Db** 130 ARNAKDCWYSYHYPRMSQVLPDPGTWNVEVFETFKNGKSWGSWFDFHVKGWNIIRDYQI 189  
| : | | | | | | | | | | | | | | | | | |

**Qy** 192 LFLKYEDMHRLDVTMVBLQARFLGVS-----CDKAQLEALTEHCHOLVDQCCEALP-- 244  
| | | | | | | | | | | | | | | | | | | | | | | |  
**Db** 190 LFLFYEDVKRDPKREIKQVMQFMGNLDEEVDVKIVLETSEKMK--NPMWNRSTVPKS 247  
| | | | | | | | | | | | | | | | | | | | | | | |

**Qy** 245 -----VGRGRVGLWKOIFTVSMNEKFDLVYKQMGKCDLTF 280  
| : | | | | | | | | | | | | | | | | | |  
**Db** 248 VLDOISPFWRKGTVGWDQKNHTFAQNDREFDIYKQMGGTSLNF 292  
| : | | | | | | | | | | | | | | | | | |

**RESULT 4**

**S10329**  
aryl sulfotransferase (EC 2.8.2.1) IV - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C;Accession: S10329; A44932  
R;Ozawa, S.; Nagata, K.; Gong, D.; Yamazoe, Y.; Kato, R.  
Nucleic Acids Res. 18, 4001, 1990  
A;Title: Nucleotide sequence of a full-length cDNA (PST-1) for aryl sulfotransferase from  
A;Reference number: S10329; MUID:90326537; PMID:2374726  
A;Accession: S10329  
A;Molecule type: mRNA  
A;Residues: 1-291 <OZA>  
A;Cross-references: UNIPROT:P17988; UNIPARC:UIP0000000B75; EMBL:X52883; NID:955760; PIDN:  
R;Yerokun, T.; Etheredge, J.L.; Norton, T.R.; Carter, H.A.; Chung, K.H.; Birckbichler, P.  
Cancer Res. 52, 4779-4786, 1992  
A;Title: Characterization of a complementary DNA for rat liver aryl sulfotransferase IV  
used hepatocarcinogenesis.  
A;Reference number: A44932; MUID:92379794; PMID:1511441  
A;Accession: A44932  
A;Status: preliminary

A:Molecule type: mRNA; protein  
A:Residues: 7-291 <YER>  
A:Cross-references: UNIPARC:UPI000017088C; GB:X68640; GB:S42994; NID:G55764; PIDN:CAA488  
A:Experimental source: liver  
A>Note: sequence extracted from NCBI backbone (NCBIN:112171, NCBIP:112178)  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: sulfoltransferase

Query Match      29.8%; Score 456; DB 2; Length 295;  
Best Local Similarity    33.7%; Pred.No.8.6e-32;  
Matches     94; Conservative    58; Mismatches 101; Indels    26; Gaps    4;

Qy    24   GVRLPPFCRGKMEETANFPVRSDVMIVTPYPSGTSLLOEVVYLVSOGADPEIGLNMID 83  
Db    13   GIPLIKYPFAETIGPQLONTAWPDLLISTPYPSGGTTWMSSEILDMIVYGGLCKGRAPIF 72

Qy    84   BOLPVLEVPQP-----GLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVITYMARNPOLDV 139  
Db    73   ARVPFLEPKCGVPGSGLETFTPTAPRLKLTHPLLSLLQSLLDOKVKVIYIARNKDVV 132

Qy    140   VSYIQFHRSRLTMSYRGTFQEFRCRRFMNDKLGYSWFHEHVQEWEHRMDSNVLFKYEDM 199  
Db    133   VSYYNFYNMAKHLDPDGTDWSPLENFMDCGEVSGYQHVKEWELRHHTHPVLVLYFYEDI 192

Qy    200   HRDLVTWVEQLARFLGVSCDKAQALTEHC--HQLVDDQC-CNAEALP-----V 245  
Db    193   KENPKREIKLIKLEFLGRSLPETVDPSIVHHTSFKKMKENCMTNYTIPTTEIMDHNVSPFM 252

Qy    246   GRGRVGWLWKDIFTVSMNEKFDELIVYKQMGKCIDLTF 280  
Db    253   RKGTTGDWNKTFTVAQNRFDAHYAKMTDCCFKP 287

RESULT 5  
S52791  
aryl sulfotransferase (EC 2.8.2.1) - human  
C:Species: Homo sapiens (man)  
C>Date: 19-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 18-Jun-1999  
C:Accession: S52791  
C:Ozawa, S.; Nagata, K.; Shinada, M.; Ueda, M.; Teuzuki, T.; Yamazoe, Y.; Kato, R.  
submitted to the EMBL Data Library, March 1994  
A:Description: Primary structures and properties of two related aryl sulfotransferases  
A:Reference number: S52791  
A:Accession: S52791  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-295 <OZA>  
A:Cross-references: UNIPARC:UPI000016B06E; EMBL:X78282; NID:g758594; PIDN:CAAS5088.1; PI:PI  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: sulfoltransferase

Query Match      29.8%; Score 456; DB 2; Length 295;  
Best Local Similarity    33.7%; Pred.No.8.6e-32;  
Matches     94; Conservative    58; Mismatches 101; Indels    26; Gaps    4;

Qy    24   GVRLPPFCRGKMEETANFPVRSDVMIVTPYPSGTSLLOEVVYLVSOGADPEIGLNMID 83  
Db    17   GVLLIKYPFAEALGPLSQSFQARDPLLSTPYPSGGTTWVSQILDWIYQGDLCKCHRAPIF 76

Qy    84   BOLPVLEVPQP-----GLDIIKELTSPRIKSHLPYRFLPSDLHNGDSKVITYMARNPOLDV 139  
Db    77   MRVPFLEPKCVGPISGMETLKNTPTAPRLKLTHPLALLPQTLLDQVKVIYVARNAKDVA 136

Qy    140   VSYIQFHRSRLTMSYRGTFQEFRCRRFMNDKLGYSWFHEHVQEWEHRMDSNVLFKYEDM 199  
Db    137   VSYYHFNMAKVIYPHPGTWESFLEXFMAGEVSGYQHVKEWELSRHTHPVLVLYFYEDM 196

Qy    200   HRDLVTWVEQLARFLGVSCDKAQALTEHC-----HQLVDDQCNAEA 242  
Db    197   KENPKREIKLIKLEFVGRSLPETVDLMVEHTSFKEMKNPMNTYTTRREFMDHSIS--- 253

Qy    243   LP-VGRGRVGIWLKIIFTVSMNEKFDLVYKQMGKCIDLTF 280

Db 254 -PFMRKGMAGDWKTTFTTVAQNERFDADYAEKMGCSLSF 291

RESULT 6  
JC5249  
aryl sulfotransferase (EC 2.8.2.1) 2 - human  
N:Alternate names: phenol sulfotransferase; sulfokinase  
C:Species: Homo sapiens (man)  
C:Date: 13-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C:Accession: JC5249  
R:Dooley, T.P.; Huang, Z.  
Biochem. Biophys. Res. Commun. 228, 134-140, 1996  
A:Title: Genomic organization and DNA sequences of two human phenol sulfotransferase genes  
A:Reference number: JC5248; MUID:97069665; PMID:8912648  
A:Accession: JC5249  
A:Molecule type: mRNA  
A:Residues: 1-295 <DOO>  
A:Cross-references: UNIPROT:P50226; UNIPARC:UPI0000136229; GB:U76619; NID:g1666920; PIDN:  
C:Comment: This enzyme catalyzes the transfer of the sulfonate group from phosphoadenosine  
C:Genetics:  
A:Gene: stp2  
A:Map position: 16p12.1-pl1.2  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: sulfotransferase

Query Match 29.8%; Score 456; DB 2; Length 295;  
Best Local Similarity 33.7%; Pred. No. 8.6e-32;  
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;

Qy 24 GVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLQEVVYLVSQGADPDEIGLMNID 83  
Db 17 GVPLIKYFAEALGPLQSFQARPDLLISTYPSKGTWVSQILDMYQGGDLEKCHRAPIF 76

Qy 84 EQLPVLEYPOP-GLDIIKELTSPLIKSHLPYRFLPSDLHNGDSKVIYMARNPXDLV 139  
Db 77 MRVPFLFKVPGIPSGMETLKNTPAPRLKTHLPLALLPQTLDDQKVKVTVARNAKDVA 136

Qy 140 VSYVQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVEHMDSNVLFKYEDM 199  
Db 137 VSYHYFHYMAKVYPHPTWESFLEKFMAGEVSGYSWYQHVQEWELSRTHPVLVLFYEDM 196

Qy 200 HRDLVTWVEQLARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAEA 242  
Db 197 KENPKREIQKILEFVGRSLPBEETVDLMVEHTSFPEMKKNPMTNVTYVRRPFMDHSIS--- 253

Qy 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280  
Db 254 -PFMRKGMAGDWKTTFTTVAQNERFDADYAEKMGCSLSF 291

RESULT 7  
G01843  
aryl sulfotransferase - human  
C:Species: Homo sapiens (man)  
C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 09-Jul-2004  
C:Accession: G01843  
R:Zhu, X.  
submitted to the EMBL Data Library, May 1995  
A:Reference number: G08606  
A:Accession: G01843  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-295 <ZHU>  
A:Cross-references: UNIPROT:P50226; UNIPARC:UPI0000154B6D; EMBL:U28169; NID:g881502; PID  
C:Superfamily: alcohol sulfotransferase

Query Match 29.8%; Score 456; DB 2; Length 295;  
Best Local Similarity 33.7%; Pred. No. 8.6e-32;  
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;

Qy 24 GVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLQEVVYLVSQGADPDEIGLMNID 83  
Db 17 GVPLIKYFAEALGPLQSFQARPDLLISTYPSKGTWVSQILDMYQGGDLEKCHRAPIF 76

Qy 84 EQLPVLEYPOP-GLDIIKELTSPLIKSHLPYRFLPSDLHNGDSKVIYMARNPXDLV 139  
Db 77 MRVPFLFKVPGIPSGMETLKNTPAPRLKTHLPLALLPQTLDDQKVKVTVARNAKDVA 136

Qy 140 VSYVQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVEHMDSNVLFKYEDM 199  
Db 137 VSYHYFHYMAKVYPHPTWESFLEKFMAGEVSGYSWYQHVQEWELSRTHPVLVLFYEDM 196

Qy 200 HRDLVTWVEQLARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAEA 242  
Db 197 KENPKREIQKILEFVGRSLPBEETVDLMVEHTSFPEMKKNPMTNVTYVRRPFMDHSIS--- 253

Qy 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280  
Db 254 -PFMRKGMAGDWKTTFTTVAQNERFDADYAEKMGCSLSF 291

RESULT 8  
157945  
phenol-sulfating phenol sulfotransferase - human  
C:Species: Homo sapiens (man)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: I57945  
R:Wilborn, T.W.; Comer, K.A.; Dooley, T.P.; Reardon, I.M.; Heinrikson, R.L.; Falany, C.N.  
Mol. Pharmacol. 43, 70-77, 1993  
A:Title: Sequence analysis and expression of the cDNA for the phenol-sulfating form of h  
A:Reference number: I57945; MUID:93140712; PMID:8423770  
A:Accession: I57945  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-295 <RES>  
A:Cross-references: UNIPROT:P50225; UNIPARC:UPI000016AEF7; GB:L19999; NID:g307342; PIDN:  
C:Superfamily: alcohol sulfotransferase

Query Match 29.7%; Score 454; DB 2; Length 295;  
Best Local Similarity 34.2%; Pred. No. 1.3e-31;  
Matches 94; Conservative 56; Mismatches 107; Indels 18; Gaps 3;

Qy 24 GVRLLPPFCRGKMEETANFPVRPSDVWIVTPKSGTSLQEVVYLVSQGADPDEIGLMNID 83  
Db 17 GVPLIKYFAEALGPLQSFQARPDLLISTYPSKGTWVSQILDMYQGGDLEKCHRAPIF 76

Qy 84 EQLPVLEYPOP-GLDIIKELTSPLIKSHLPYRFLPSDLHNGDSKVIYMARNPXDLV 139  
Db 77 MRVPFLFKVPGIPSGMETLKNTPAPRLKTHLPLALLPQTLDDQKVKVTVARNAKDVA 136

Qy 140 VSYVQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVEHMDSNVLFKYEDM 199  
Db 137 VSYHYFHYMAKVYPHPTWESFLEKFMAGEVSGYSWYQHVQEWELSRTHPVLVLFYEDM 196

Qy 200 HRDLVTWVEQLARFLGVSCDKAQLEALTEHC-----HQLVDQCCNAEA 245  
Db 197 KENPKREIQKILEFVGRSLPBEETVDLMVEHTSFPEMKKNPMTNVTYVRRPFMDHSISPPFM 256

Qy 246 GRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280  
Db 257 RKGMDAGDWKTTFTTVAQNERFDADYAEKMGCSLSF 291

RESULT 9  
S52794  
aryl sulfotransferase (EC 2.8.2.1) - human  
C:Species: Homo sapiens (man)  
C:Date: 19-May-1995 #sequence\_revision 03-Aug-1995 #text\_change 09-Jul-2004  
C:Accession: S52794  
R:Ozawa, S.; Nagata, K.; Shimada, M.; Ueda, M.; Tsuzuki, T.; Yamazoe, Y.; Kato, R.  
submitted to the EMBL Data Library, March 1994  
A:Description: Primary structures and properties of two related aryl sulfotrans-ferases  
A:Reference number: S52791  
A:Accession: S52794  
A:Status: preliminary  
A:Molecule type: mRNA







A;Title: Oestrogen sulfotransferase: molecular cloning and sequencing of cDNA for the bo  
A;Reference number: S29045; MUID:90147277; PMID:32711383  
A;Accession: S29045  
A;Molecule type: mRNA  
A;Residues: 1-295 <NAS>  
A;Cross-references: UNIPROT:P19217; UNIPARC:UPI00001136223; EMBL:M54942; NID:g163462; PID  
A;Experimental source: placenta  
R;Adams, J.B.  
Biochim. Biophys. Acta 1076, 282-288, 1991  
A;Title: Enzymic synthesis of steroid sulphates XVII. On the structure of bovine estroge  
A;Reference number: S13486; MUID:91152101; PMID:1900200  
A;Accession: S18222  
A;Molecule type: protein  
A;Residues: 146-160;206-220;257-259 <ADA>  
A;Cross-references: UNIPARC:UPI0000175826; UNIPARC:UPI0000175827; UNIPARC:UPI0000175828  
A;Experimental source: placenta  
C;Complex: homodimer  
C;Superfamily: alcohol sulfotransferase  
C;Keywords: homodimer; phosphoprotein; sulfotransferase

Query Match 29.2%; Score 446; DB 2; Length 295;  
Best Local Similarity 34.2%; Pred. No. 6.3e-31;  
Matches 100; Conservative 57; Mismatches 111; Indels 24; Gaps 5;

QY 10 STFGEPESKYFBHGVRLLPFCRGKMEIAN---FPVRPSDVWIVTPKSGTSLQEVVY 66  
DB 3 SSKPSPSDYFGKLGGI---PMVKFIEQFHNVEEFAPRDDLVIVTPKSGTTLSEITC 59  
QY 67 LVSQGADPDEIGLMTDEQLPVL---EYPQGLDIIKELTSPRLIKSHLPYRFLPSDLH 122  
DB 60 MYNNGDVSKCEDVIFNRPVPLECSTEHVMGVKQNLNEMASPRIVKSHLPVKLLPVSFW 119  
QY 123 NGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFFHVOBF 182  
DB 120 ERNKIIYLSRAKDVVSVYFLLMVTAPDPDSQDFVEKMDGEVPGYGSFETHKSW 179  
QY 183 WEHRMDSNVFLKYEDMDHRDLVTMVEQLARFLGVSCDKAQLAEALTEHCH- 231  
DB 180 WEKSKNPQVFLFYEDMDKENIRKVMKLEFLGRKASDELVDKIIKHTSFQEMKNPSTN 239  
QY 232 --QLVDQCCNAEALP-VGRGRVGLWKDIETVSNNEKFDLVYKQKMGKCDLTF 280  
DB 240 YTTLPDEVNQVSPFMRKGVGDWKNHFTVALNEKFDHMYEQQMKGSTLKF 291

RESULT 16  
S52399  
aryl sulfotransferase (EC 2.8.2.1) - human  
C;Species: Homo sapiens (man)  
C;Date: 08-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 18-Jun-1999  
C;Accession: S52399  
R;Jones, A.L.; Hagen, M.; Coughtrie, M.W.H.; Roberts, R.C.; Glatt, H.R.  
submitted to the EMBL Data Library, February 1995  
A;Description: Human platelet phenol:sulfotransferases: cDNA cloning, stable expression i  
A;Reference number: S52398  
A;Accession: S52399  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-295 <JON>  
A;Cross-references: UNIPARC:UPI000016B4D7; EMBL:X84654; NID:g671641; PIDN:CAA59147.1; PI  
A;Superfamily: alcohol sulfotransferase  
C;Keywords: sulfotransferase

Query Match 29.0%; Score 444; DB 2; Length 295;  
Best Local Similarity 33.5%; Pred. No. 9.4e-31;  
Matches 92; Conservative 57; Mismatches 108; Indels 18; Gaps 3;

QY 24 GVRLLPFCRGKMEIANFPVRPSDVWIVTPKSGTSLQEVVYLVSQGADPDEIGLMTD 83  
DB 17 GVPLIKYFAELGPIQSFQARPDLLISTYPKSGTTWVSQILDMIYQGGDEKCHRAPIF 76  
QY 84 EQLPVLVEYFPQ-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYARNPKDLV 139

DB 77 MRVPLEFKARQIPSGMETLKDTPAPRLKTHLPALLPQTLDDOKVYVYVARNAKDVA 136  
QY 140 VSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFFHVOBFWEHRMDSNVFLKYEDM 199  
DB 137 VSYHYFHTMKVHPEPGTWSDFLEKFMVGEVSGYSWYQHVQEWMLGSLRTHPVLVLYFYEDM 196  
QY 200 HRDLVTMVEQLARFLGVSCDKAQLAEALTEHC---HQLVDQCCNAEALP-----V 245  
DB 197 KENPKREIQKLEFVGHSLPEETVDFMVQHTSFKEMKKNPMNTYTVFQEFMDHSISPFM 256  
QY 246 GRGRVGLWKDIETVSNNEKFDLVYKQKMGKCDLTF 280  
DB 257 RKGMDGDKTFTTVAQNERFDADYAEKMGAGCSLSF 291

RESULT 17  
JC5885  
thyroid hormone sulfotransferase (EC 2.8.2.-) B2 - human  
C;Species: Homo sapiens (man)  
C;Date: 11-Mar-1998 #sequence\_revision 11-Mar-1998 #text\_change 09-Jul-2004  
C;Accession: JC5885  
R;Fujita, K.; Nagata, K.; Ozawa, S.; Sasano, H.; Yamazoe, Y.  
J. Biochem. 122, 1052-1061, 1997  
A;Title: Molecular cloning and characterization of rat ST1B1 and human ST1B2 cDNAs, enc  
A;Reference number: JC5884; MUID:98104061; PMID:9443824  
A;Accession: JC5885  
A;Molecule type: mRNA  
A;Residues: 1-296 <FUU>  
A;Cross-references: UNIPROT:O43704; UNIPARC:UPI0000073458; DBJ:D89479; NID:g2826145; PI  
A;Experimental source: liver  
C;Comment: This enzyme mediates the transfer of ISO3- from 3'-phosphoadenosine 5'-phosph  
C;Superfamily: alcohol sulfotransferase  
C;Keywords: sulfotransferase

Query Match 28.8%; Score 440.5; DB 2; Length 296;  
Best Local Similarity 33.9%; Pred. No. 1.9e-30;  
Matches 94; Conservative 54; Mismatches 110; Indels 19; Gaps 3;

QY 23 HGVRLPPFCRGKMEIANFPVRPSDVWIVTPKSGTSLQEVVYLVSQGADPDEIGLMTI 82  
DB 16 HGYPMTCAFASNWEKIEQFHSRPDDIVATVPKSGTTWVSEIIDMLNDGIEKCKRGFI 75  
QY 83 DEQLPVLEYPOPL-----DIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYARNPKD 137  
DB 76 TEKVPMEMLTLPGLRTSGIEQLEKNPSRIVKTHLPTDLLPKSPWENNCCKMIYLARNAD 135  
QY 138 LVSVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFFHVOBFWEHRMDSNVFLKYE 197  
DB 136 VSVSYHFDLMNINLPQFPFTWEEYLEKFLTKGVAYGSWFTVHKVNMWKKKGHPILFLYYE 195  
QY 198 DMHRDLVTMVEQLARFLGVSCDKAQLAEALTEHCHOLV---DQCCNAEALP----- 244  
DB 196 DMKENPKREIKKIIIPLEKNLDELDIRIIHTSEVKNKDNPLVNYTHLPTTVMDSKSP 255  
QY 245 -VGRGRVGLWKDIETVSNNEKFDLVYKQKMGKCDLTF 280  
DB 256 FMRKGTAGDKNRYFTVAQNEKFDALYETEMSKTALQF 292

RESULT 18  
I73679  
estrogen sulfotransferase isoform 6 - rat  
C;Species: Rattus sp. (rat)  
C;Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 18-Jun-1999  
C;Accession: I73679  
R;Falany, J.L.; Krasnykh, V.; Mikhcheva, G.; Falany, C.N.  
J. Steroid Biochem. Mol. Biol. 52, 35-44, 1995  
A;Title: Isolation and expression of an isoform of rat estrogen sulfotransferase.  
A;Reference number: I56606; MUID:95161323; PMID:7857871  
A;Accession: I73679  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-295 <RES>

A; Cross-references: UNIPARC:UPI00013622; GB:S76490; NID:g913356; PIDN:AAB33442.1; PID:13622  
C; Superfamily: alcohol sulfotransferase

Query Match 28.4%; Score 435; DB 2; Length 295;  
Best Local Similarity 31.8%; Pred. No. 5.6e-30;  
Matches 93; Conservative 62; Mismatches 105; Indels 32; Gaps 4;

Qy	14	EFESKYEFPHGVRUPPFCRGKWEELANFPVPSPDVWIVTPYKSGTSLQEVVYVSQAD	73
Db	7	EYEVFGDHGVLVDLFTKYWEDVETFSARPDDLVLVTPYKSGSTWTGIVDMTIEKGD	66
Qy	74	PDEIGLMMIDQLPVLEYQPQ-----GLDIIKELTSPLRIKSHLPYRFLPSPDLHNGDSKVI	129
Db	67	VEKCKEDAIFNRIPLYECRNEDLNGITKQKEKESPRIVKTHLPKALJPASPWEKNCXII	126
Qy	130	YMAENPKDLVVSYTYQFHRSLRTMSYRGTFQBFCCRFRMNDKLGYSWGFHVFQFWEHRMDS	189
Db	127	YLCRNADVVSYYTYFFLLIKSYNPKSFSFEVEKFMEGVPYGSWYDHVKSWEKSNKS	186
Qy	190	NVLFLKYEDMHRDLVTWVQELARFLGVSCDKAQALTEHCHQLV-----DQCNC	239
Db	187	RVLFMFVEDMKEDTRREVVKLIETL-----ERDPLAELVDKIIQHTSFQBMKNPCFN	239

RESULT 19

JC2229

estrogen sulfotransferase (BC 2.8.2.-) - human

C:Species: Homo sapiens (man)

C:Date: 28-Aug-1985 #sequence revision 07-Oct-1994 #text\_change 09-Jul-2004

C:Accession: JC2229; I38249; I56607

R:Aksay, I.A.; Wood, T.C.; Weinshilboum, R.

Biochem. Biophys. Res. Commun. 200, 1621-1629, 1994

A:Title: Human liver estrogen sulfotransferase: identification by cDNA cloning and expression

A:Reference number: JC2229; MUID:94242031; PMID:8185618

A:Accession: JC2229

A:Molecule type: mRNA

A:Residues: 1-294 <AKS>

A:Cross-references: UNIPROT:P49888; UNIPARC:UPI00000623C8; GB:U08098; NID:G488282; PIDN:

A:Experimental source: liver

R:Her, C.; Aksay, I.A.; Kimura, S.; Brandriff, B.F.; Wasmuth, J.J.; Weinshilboum, R.M.

Genomics 29, 16-23, 1995

A:Title: Human estrogen sulfotransferase gene (STE): Cloning, structure and chromosomal

A:Reference number: I38249; MUID:96079087; PMID:8530066

A:Residues: 1-294 <RES>  
A:Cross-references: UNIPARC:UPT00000623C8; EMBL:U20521; NID:g1079511; PIDN:AAC50286.1; R:Falany, C.N.; Krasnykh, V.; Falany, J.L.  
J. Steroid Biochem. Mol. Biol. 52, 529-539, 1995  
A:Title: Bacterial expression and characterization of a cDNA for human liver estrogen su  
A:Reference number: I56607; MUID:95298714; PMID:7779757  
A:Accession: I56607  
A:Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: mRNA  
A:Residues: 1-294 <RES>  
A:Cross-references: UNIPARC:UPT00000623C8; GB:S77383; NID:g998887; PIDN:AAB344601.1; PID:  
C:Comment: This enzyme catalyzes the sulfate conjugation of estrogen.

C:Gene: GDB:STE  
A:Gene: GDB:STE  
A:Cross-references: GDB:342104; OMIM:600043  
A:Map position: 16pter-16qter  
A:Introns: 49/1; 91/1; 123/3; 166/1; 197/3; 258/1  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: sulfotransferase

Query Match 28.4%; Score 434; DB 2; Length 294;  
Best Local Similarity 33.3%; Pred. No. 6.8e-30;  
Matches 95; Conservative 56; Mismatches 116; Indels 18;  
Gaps 3;

Qy	14	EPESKYFEFHGVRUPPPCGRKHESIANPPVRPSDVITVTPKSGTSLLRVVTVLVSQGAD	73
Db	6	DYIEKFEFEEHILMYKDFVKVYDWNVEAFQARPDOLVIATVPKSGTTWSSIVVMYKEGD	65
Qy	74	PBEIGLMIWIDEOLPVLVEYPOB---GLDITKELTSPLRIKSHLPYRFLPSDLHNGSKVI	129
Db	66	VEKKCEDVIFNRIFFLECRKENLMNGVKQIDEMNSPRIKVTHLPPELLPASFWKDOCKII	125
Qy	130	YMARNPKDLVSYQYOFHRSLSLMTSYRTQTOFCRRFMNDKLGYSWFHVEHQEPWEHRMDS	189
Db	126	YLICRAKADVAVSYFYFFLWAGHPNPGSPFPEFVEKFMQGVPGYKWKVYKHWKSWWEKGP	185
Qy	190	NVLFLUKYEDMHRDLVTWYSQLARFLGVSCDKAQLEALTEHCH-----QLVDQ	236
Db	186	RVLFLFYEDLKDRTKEVILKILHFERKPSBELVDRIIHHTSFOEMKNPNSTNYTTLDPD	245
Qy	237	CCNBAALP-VGRGRVGLWKDIFTYSNMNEKFDLVYKQRMGKCDLTF	280
Db	246	INNOKLSPFMKRGITGDWKNHFTVALNEKEDKHVEOOMKESTLKF	290

## RESULT 20

JE0186  
sulfotransferase (EC 2.8.--.) SULT1B1 - mouse  
C:Species: Mus musculus (house mouse)  
C:date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change  
C:Accession: JE0186  
R:Saeki, Y.; Sakakibara, Y.; Araki, Y.; Yanagisawa, K.; Suiko, I.  
J. Biochem. 124, 55-64, 1998  
A:Title: Molecular cloning, expression, and characterization of  
A:Reference number: JE0186; PMID:98303980; CUID:9644246

Query Match	28.3%	Score 432.5	DB 2	Length 299
Best Local Similarity	35.1%	Pred. No. 9.4e-30		
Matches 94	Conservative 50	Mismatches 97	Indels 27	Gaps 4

36	QY	EEIANFPVRPSDWIVVYPKSCHSLLOEVYVLSQGNADPDEICLMMIDQLPVLVYPPQGS	95
29	Db	ERIEEFQSTPGDIVITVYPSGTTWSEIVDMVLNPGNVKECKRDRVITSKVPMLLESLVSG	88
96	QY	L-----DIIKELTSPRIKSHLPVRFLPSDLHNGDGSKIYMARNPDXLVVSYVYQFHRSLR	150
89	Db	IRISGVELLKKTSPRIIKTHLPIDLLPKSFWEKCKMVIYLAENGKDVASYVYHFDLMNS	148
151	QY	TMSYRGTTQETCRPMNDKLYGSGWPEHQVFEHHRMDSNVLFKYEDMHRDLVTWVEQL	210
149	Db	INPLPGTWEEYLEKFLAGNVAYGSGWFHVSWMKEEHPHLLYLYEELKQNPKKEIKKI	208
211	QY	ARFLGVSCDQAKLALTEHC-----HQLVDQCCNAEALPVG-----RGRVGL	252
209	Db	ASFLDKTILDEALDRIVHTHSFEMKENKPLV----NYTHLPTAMDHSSKSPWPKGI	264
253	QY	WKDIFTVSMNEKFDLVYKQKMGKCDLTF	280
265	Db	WNGYFTMTQTQGFDAVYKQKMSGTTLEF	292

## RESULT 21

```

testis-specific estrogen sulfotransferase - mouse
I53296
RESOURC 21
C:Species: Mus sp. (mouse)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 18-Jun-1999
C:Accession: I53296
R:Song, W.C.; Moore, R.; McLachlan, J.A.; Negishi, M.
Endocrinology 136, 2477-2484, 1995

```



Qy 81 NIDEQLPVLEV--POP---GLDIIKELTSPRLIKSHLPVRFPLPSDLHNGDSKVIYMARNP 135  
Db 82 NYDRHPFIEWTLPSPLNSGLDLANKMPSPTLKTHLPVHMLPPSPFWKENSKIIVARNNA 141  
Qy 136 KDLVVSYYQPHRSLSRTMSYRGTFQBFRCRRFMNDKLGYSWFEHVSQFWEHRMDSNVLFJK 195  
Db 142 KDLVVSYYVFRSMNKLDPDGTGLGYSIQFQKAGKVLGWSYDVKWGMVDKQDQHRILYLF 201  
Qy 196 YEDMRDLVTWVEQLARFLGVSQCDKAQLEALTEHCHQLV---DQCCNAELP----- 244  
Db 202 YEDMKEDPKREIKTKIAKLEKDISSEVLNKIYHTSFQVMKNPMMANYTTLPSSIMDRHSI 261  
Qy 245 ---VGRGVGLWKDIFTVSMNEKFDLVVKQKMGKCDLTF 280  
Db 262 SFFMRKGMFGDKWKNVFTVAQSDDFEDIRRMAGSNIITF 300

RESULT 25  
JF0197  
phenol sulfotransferase (EC 2.8.--) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 09-Jul-2004  
C:Accession: JF0197; JG0188  
R:Sakakibara, Y.; Yanagisawa, K.; Takami, Y.; Nakayama, T.; Suiko, M.; Liu, M.C.  
Biochem. Biophys. Res. Commun. 247, 681-686, 1998  
A:Title: Molecular cloning, expression, and functional characterization of novel mouse  
A:Reference number: JG0196; MUID:98321187; PMID:9647753  
A:Accession: JF0197  
A:Molecule type: mRNA  
A:Residues: 1-295 <SAK>  
A:Cross-references: UNIPROT:Q35401; UNIPARC:UPI00000566C3; GB:AF026073; NID:g2570899; P1  
R:Liu, M.; Sakakibara, Y.; Liu, C.  
Biochem. Biophys. Res. Commun. 254, 65-69, 1999  
A:Title: Bacterial expression, purification, and characterization of a novel mouse sulfo  
A:Reference number: JG0188; MUID:99121190; PMID:9920733  
A:Accession: JG0188  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-295 <LIU>  
A:Cross-references: UNIPARC:UPI00000566C3; GB:AF026073  
A:Comment: This enzyme catalyzes the transfer of a sulfonate group from the active sulfa  
P.  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: transferase

Query Match 27.0%; Score 413; DB 2; Length 295;  
Best Local Similarity 31.7%; Pred. No. 4.5e-28;  
Matches 83; Conservative 63; Mismatches 98; Indels 18; Gaps 2;

Qy 37 ETANFPVPSDVWIVTYPKSGTSLQEVVLYVSQGDADPEIGLNMIDQLPVLEYQPQGL 96  
Db 30 QVESPEARDDLLISTYPKSGTGWVSEILDLLYNNGDAEKCKRDAIYKRVPMELIIPGI 89  
Qy 97 ----DIKELTSPRLIKSHLPVRFPLPSDLHNGDSKVIYMARNPDLVVSYYQPHRSLSRTM 152  
Db 90 TNGVEMLNMPSPRIVKTHLPVQLLPSSFWKNDCKLIIVARNAKDVVVSYYFYQWAKIH 149  
Qy 153 SYRGTFQBFRCRRFMNDKLGYSWFEHVSQFWEHRMDSNVLFJKYEDMRDLVTWVEQLAR 212  
Db 150 PEPGTWEEFLEKFMAGQVSFGPWYDVKWSWNEKREYRILYLFYEDMKENPKCEIQILK 209  
Qy 213 FLGVSQCDKAQLEALTEHCH-----QLVDQCCNAELPVGGRVGLWKDIFT 258  
Db 210 YLEKIDPEILNKILYHSFSVMKNPNSANYTTMMKEEMDHVSFPMRKRGISGDWKNQFT 269  
Qy 259 VSMNEKFDLVVKQKMGKCDLTF 280  
Db 270 VAQYEKFEEDYVKKMEDSTLKP 291

RESULT 26  
JG5000  
aryl sulfotransferase (EC 2.8.2.1), phenol-preferring - bovine

C:Species: Bos primigenius taurus (cattle)  
C:Date: 13-Dec-1996 #sequence\_revision 21-Jan-1997 #text\_change 09-Jul-2004  
C:Accession: JG5000; S59269; S59070  
R:Henry, T.; Kliewer, B.; Palmatier, R.; Ulphani, J.S.; Beckmann, J.D.  
Gene 174, 221-224, 1996  
A:Title: Isolation and characterization of a bovine gene encoding phenol sulfotransferase  
A:Reference number: JG5000; MUID:97045816; PMID:8890738  
A:Accession: JG5000  
A:Molecule type: DNA  
A:Residues: 1-294 <HEN>  
A:Cross-references: UNIPROT:P50227; UNIPARC:UPI000013622A; GB:U34753; NID:g1276632; PIDN  
R:Schaus, S.J.; Henry, T.; Palmatier, R.; Halvorson, L.; Dannenbring, R.; Beckmann, J.D.  
Biochem. J. 311, 209-217, 1995  
A:Title: Characterization of bovine tracheobronchial phenol sulphotransferase cDNA and  
A:Reference number: S59070; MUID:96003918; PMID:7575456  
A:Accession: S59269  
A:Molecule type: mRNA  
A:Residues: 1-294 <SCH>  
A:Cross-references: UNIPARC:UPI000013622A; GB:U35253; NID:g1006836; PIDN:AAA85510.1; PID  
A:Accession: S59070  
A:Molecule type: protein  
A:Residues: 23-36738-47; 79-96; 134-144; 207-226; 234-248 <SCW>  
A:Cross-references: UNIPARC:UPI0000175829; UNIPARC:UPI000017582A; UNIPARC:UPI000017582B;  
C:Comment: This enzyme belongs to a family of cytosolic enzymes that catalyze the transf  
C:Genetics:  
A:Gene: pst  
A:Introns: 50/1; 93/3; 124/3; 167/1; 198/3; 259/1  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: sulfotransferase

Query Match 26.9%; Score 412; DB 2; Length 294;  
Best Local Similarity 33.1%; Pred. No. 5.4e-28;  
Matches 91; Conservative 56; Mismatches 108; Indels 20; Gaps 4;

Qy 17 SKYFEFGHVRLLPFCRCGMEEIANFPVRPSDVWIVTYPKSGTSLQEVVLYVSQGDADPE 76  
Db 12 AKYVK--GIPLIKYFAEALGPLSEFAWPDLLISTYPKSGTTWVSEILDLIYQEGDLEK 69  
Qy 77 IGLNMIDQLPVLEYQPQ---GLDIIKELTSPRLIKSHLPVRFPLPSDLHNGDSKVIYMA 132  
Db 70 CORAPVFLVRFLVFEFSAPGVPTGVLELLKDTTPAPRLKTHLPALLPKTLDDQKVKVIYA 129  
Qy 133 RNPKDLVVSYYQPHRSLSRTMSYRGTFQBFRCRRFMNDKLGYSWFEHVSQFWEHRMDSNVL 192  
Db 130 RNAKDVAVSYYHFRMAKVHPDPTGWSFLEKFMAGEVCYGSWYQHWELSHSTHVPVL 189  
Qy 193 FLKYEDMRDLVTWVEQLARFLGVSQCDKAQLEAL---TEHCHQLVDQCCNAELPVG--- 246  
Db 190 YLFYEDIKEDPKREIQKILEFGRSLPEETVDHIVQRTSPFKMKKNPMTNYSTIPTAYMD 249  
Qy 247 -----RGRVGLWKDIFTVSMNEKFDLVYKQKM 273  
Db 250 HSISAFMRKGITGDMKSTFTVAQNELFEAHYAKKM 284

RESULT 27  
A44011  
adrenocortical estrogen sulfotransferase - guinea pig  
C:Species: Cavia porcellus (guinea pig)  
C:Date: 27-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A44011  
R:Oeda, T.; Lee, Y.C.; Driscoll, W.J.; Chen, H.C.; Strott, C.A.  
Mol. Endocrinol. 6, 1216-1226, 1992  
A:Title: Molecular cloning and expression of a full-length complementary DNA encoding th  
A:Reference number: A44011; MUID:93024479; PMID:1406700  
A:Accession: A44011  
A>Status: preliminary  
A:Molecule type: mRNA; protein  
A:Residues: 1-296 <OED>  
A:Cross-references: UNIPROT:P49887; UNIPARC:UPI000013622A; GB:U09552; GB:S45979; NID:g49  
A:Note: sequence extracted from NCBI backbone (NCBIN:115159, NCBI:P115160)  
C:Superfamily: alcohol sulfotransferase



Db 11 SSSEKNVSEMGWFFGGYFRVKGIPFPVGMVSPESLSAENTSN--VRDDIFIVITYPKS 68  
Qy 57 GTSLLQEVVYLVSGADPEIGLNMIDQLPVLVEYPPQGLDIIKELTSPRLIKSHLPYRF 116  
Db 69 GTNMWIEIVCLIKDGDPSWIRSEPIWQAPWCEIISAFNVL--DRPSPRIMSHLPIEL 127  
Qy 117 LPSDLHNGDSKVIYMARPKDLVSYQFHRSLRTMSYRGTFQEFRCRRMNDKLGYSWFP 176  
Db 128 FTKAPFSSKAKVIYGRNPRDVVSLYYYSKIAGQLKDPGTPDQPLQNLKGEVQFGSWF 187  
Qy 177 EHVOEFWEHRMDSNVLFLLKYEDMHRDLVTWVEQLARFLGVSCDKAQLALTEH---CHQL 233  
Db 188 DHIKGWIWMQENFLFYBELQDLRGSVQRICEFLRGELGEALSSVVAHSAFAAMK 247  
Qy 234 VDQCCNABALPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQKM 273  
Db 248 ANTMNSYLLPASLLDHRQGEFLRKGISGDMKNHFTVAQSEAFDSVYREQM 298  
RESULT 31  
A54026  
alcohol sulfotransferase (EC 2.8.2.2) - guinea pig  
C:Species: *Cavia porcellus* (guinea pig)  
C:Date: 02-Aug-1994 #sequence\_revision 02-Aug-1994 #text\_change 09-Jul-2004  
C:Accession: A54026  
R:Lee, Y.C.; Park, C.S.; Strott, C.A.  
J. Biol. Chem. 269, 15838-15845, 1994  
A:Title: Molecular cloning of a chiral-specific 3alpha-hydroxysteroid sulfotransferase.  
A:Reference number: A54026; MUID:94253178; PMID:8195238  
A:Accession: A54026  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-287 <LEE>  
A:Cross-references: UNIPROT:P50234; UNIPARC:UPI0000170743; GB:U06871; NID:G508212; PIDN:  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: sulfotransferase  
Query Match 24.6%; Score 376.5; DB 2; Length 287;  
Best Local Similarity 32.8%; Pred. No. 6.1e-25;  
Matches 88; Conservative 49; Mismatches 114; Indels 17; Gaps 5;  
Qy 22 PHGVLPFPCKGK--MEEIAN-FPYRPSDVMTVTPKSGTSLQEVVYLVSGADPDIG 78  
Db 8 FEGIRPPMVCFSPILREVRDKFLVKDEDTITVTPKSGTNWNLIEVCLILSKGDPKLVQ 67  
Qy 79 LWNIDEQLPVLEYPOGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPXDL 138  
Db 68 SVPNWDRSPWIEF-TGGYELVKGQKDPVYITSHLPFLPFPKSFSSKAKVIYCIENPRDA 126  
Qy 139 VVSYQFHRSLRTMSYRGTFQEFRCRRFMNDKLGYSWFEHQFWEHRMDSNVLPFLKYED 198  
Db 127 LVSGYFFLSKMNVTKEPTELQYMBWFLQGNVIYGSWFEHVRGWSLMSREMENVLSYED 186  
Qy 199 MHRDLVTWVEQLARFLGVSCDKAQLALTEHC---HQLVDQCCNABALP-----V 245  
Db 187 LIKTRSTVEKICQFLKGLKPEETDLVLKYSFQFMKENMSNFTLLPHAYTTGFTLL 246  
Qy 246 GRGRVGLWKDIFTVSMNEKFDLVYKQKM 273  
Db 247 RKGTVGDMKNHFTVAQAEAFDKIYQEKM 274  
RESULT 32  
JE0152  
alcohol sulfotransferase (EC 2.8.2.2) - rabbit  
N:Alternate names: ASH-RB2; hydroxysteroid sulfotransferase  
C:Species: *Oryctolagus cuniculus* (domestic rabbit)  
C:Date: 13-Jun-1998 #sequence\_revision 10-Jul-1998 #text\_change 09-Jul-2004  
A:Accession: JE0152  
R:Yoshinari, K.; Nagata, K.; Shiraga, T.; Iwaseki, K.; Hata, T.; Ogino, M.; Ueda, R.; Fu  
J. Biochem. 123, 740-746, 1998  
A:Title: Molecular cloning, expression, and enzymatic characterization of rabbit hydroxy

A:Reference number: JE0152; MUID:98207066; PMID:9538269  
A:Accession: JE0152  
A:Molecule type: mRNA  
A:Residues: 1-286 <YOS>  
A:Cross-references: UNIPROT:O62648; UNIPARC:UPI00000886AB; DBJ:AB006053; NID:G3036938;  
A:Experimental source: liver  
C:Comment: This protein shows high activities to both hydroxysteroids and amines.  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: sulfotransferase  
Query Match 24.2%; Score 371; DB 2; Length 286;  
Best Local Similarity 32.8%; Pred. No. 1.8e-24;  
Matches 90; Conservative 51; Mismatches 113; Indels 20; Gaps 6;  
Qy 15 PESKYFFHGVRLPPFCRGKMEIEANFPVRPSDVMTVTPKSGTSLQEVVYLVSGADP 74  
Db 8 YEGIAFPFVGPN--PELMARKAYEE---FVMEEDVLTVPKSGTNWNLIEILLSRSGDA 63  
Qy 75 DEIGLNMIDQLPVLEYPOGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARN 134  
Db 64 TQIOSVPIWVRSPWVE-TVSGYEDLKTMSERLLSTHLPIHFPKSIHTTKAKVIYLMRN 122  
Qy 135 PKDLVVSYYQFHRSLRTMSYRGTFQEFRCRRFMNDKLGYSWFEHQFWEHRMDSNVLF 194  
Db 123 PRDLVVSYYQFHMVYKVFVENTKSLQEYFELFDGNNVVGWFDHVGWLSLRDQKNFLL 182  
Qy 195 KYEDMHRDLVTWVEQLARFLGVSCDKAQL-----EALTEHCHQLVDQCCNABALPV 245  
Db 183 SYEELKQDTRSTIEKICHFLGKLEPEINLVLNKNSFQAMKENKMSNYLLSEVDLIE 242  
Qy 246 GRG---RVGL---WKDIFTVSMNEKFDLVYKQKM 273  
Db 243 GHGSFMRKGISGDMKNHFTVTAQAEKFDKVFKEKM 276  
RESULT 33  
I65760  
alcohol sulfotransferase (EC 2.8.2.2) - rat  
N:Alternate names: hydroxysteroid sulfotransferase; senescence marker protein 2B  
C:Species: *Rattus norvegicus* (Norway rat)  
C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 09-Jul-2004  
C:Accession: I65760; I52407  
R:Watabe, T.; Ogura, K.; Sataekawa, M.; Okuda, H.; Hiratsuka, A.  
Chem. Biol. Interact. 92, 87-105, 1994  
A:Title: Molecular cloning and functions of rat liver hydroxysteroid sulfotransferases  
A:Reference number: I52849; MUID:94306585; PMID:8033273  
A:Accession: I65760  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-285 <RES>  
A:Cross-references: UNIPROT:P50235; UNIPARC:UPI00001361D2; GB:D14989; NID:G303801; PIDN:  
R:Song, C.S.; Kim, J.M.; Roy, A.K.; Chatterjee, B.  
Biochemistry 29, 542-551, 1990  
A:Title: Structure and regulation of the senescence marker 2 gene promoter gene promoter  
A:Reference number: I52408; MUID:90148982; PMID:2302387  
A:Accession: I52407  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 2-68 <RE2>  
A:Cross-references: UNIPARC:UPI00000E82CF; GB:M29302; NID:G206995; PIDN:AAA42152.1; PID:  
C:Genetics:  
A:Intron: 46/1  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: sulfotransferase  
Query Match 23.6%; Score 361.5; DB 2; Length 285;  
Best Local Similarity 33.6%; Pred. No. 1.2e-23;  
Matches 93; Conservative 50; Mismatches 109; Indels 25; Gaps 8;  
Qy 17 SKYFEFHGVRLPPFCRGKMEIEAN----FPVRPSDVMTVTPKSGTSLQEVVYLVSGA 72  
Db 3 SDYTFEGIPFPAPWFSK-EILENSCKFKVVKVEDDILITYPKSGTNWNLIEIVCIQTGK 61

QY 73 DPEITGLMNDQLPVLVEYPOQGLDIIKELTSPRLIKSHLPVRFPLPSDLHNGDSKVIYMA 132  
Db 62 DKKIQSQMPIDWRSPWIE--TGSQYDKLTQMEGPRMLTSHLPMLHFSKSLFSSKAKVIYLI 120  
QY 133 RNPKDLVSVYYOFRSL---RTMSVYRGTFQECRFRFMDKLGYSWFEHQVFEHWRMDS 189  
Db 121 RNPRLVLSAYFWSKIALEKKPDSLGTVEW---FLKGNVAYGSWFEHQVFEHWRMDS 177  
QY 190 NVLFKYEDMHRDLVTWVEQLARFVSCDKAQLALTEH-CHQLVDQ-----CCNAE 241  
Db 178 NFLVLVYEDMKDWTGSIKKICDFLCKLEPDELNLVLKYSSFQVVKNMNSYSLMEKE 237  
QY 242 ALPVG-----RSGVLGMDIFTVSNNEKFDLVYKQKM 273  
Db 238 LILTGFTFMRKTTNDWKNHFTVAQAEAFDKVQEKM 274  
RESULT 34  
160190  
alcohol sulfotransferase (EC 2.8.2.2) a - rat  
N:Alternate names: BAST-I; bile acid sulfotransferase I; hydroxysteroid sulfotransferase  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Aug-1996 #sequence revision 02-Aug-1996 #text change 31-Dec-2004  
C:Accession: I60190; S33511; A34822; A37395; S35600; S35601; A60949; I52408  
R:Ogura, K.; Satsukawa, M.; Okuda, H.; Hiratsuka, A.; Watabe, T.  
Chem. Biol. Interact. 92, 129-144, 1994  
A:Title: Major hydroxysteroid sulfotransferase STA in rat liver cytosol may consist of b  
A:Reference number: I60190; MUID:94306558; PMID:8033248  
A:Accession: I60190  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-284 <OGU1>  
A:Cross-references: UNIPROT:P22789; UNIPROT:Q63551; UNIPARC:UPI0000170BDA; EMBL:X63410;  
R:Ogura, K.; Satsukawa, M.; Narihata, H.; Suzuki, K.; Mori, Y.; Okuda, H.; Wong, K.P.; W  
submitted to the EMBL Data Library, December 1991  
A:Reference number: S33511  
A:Accession: S33511  
A:Molecule type: mRNA  
A:Residues: 1-284 <OGU2>  
A:Cross-references: UNIPARC:UPI0000170BDA; EMBL:X63410; NID:g311661; PIDN:CAA45007.1; PI  
R:Ogura, K.; Kajita, J.; Narihata, H.; Watabe, T.; Ozawa, S.; Nagata, K.; Yamazoe, Y.; K  
Biochem. Biophys. Res. Commun. 166, 1494-1500, 1990  
A:Title: CDNA cloning of the hydroxysteroid sulfotransferase STA sharing a strong homolo  
A:Reference number: A34822; MUID:90165963; PMID:2306259  
A:Accession: A34822  
A:Molecule type: mRNA  
A:Residues: 1-155, 'Y', 157-284 <OGU3>  
A:Cross-references: UNIPARC:UPI0000161B81; GB:M33329; NID:q207076; PIDN:AAA42183.1; PID:  
R:Ogura, K.; Sohtome, T.; Sugiyama, A.; Okuda, H.; Hiratsuka, A.; Watabe, T.  
Mol. Pharmacol. 37, 848-854, 1990  
A:Title: Rat liver cytosolic hydroxysteroid sulfotransferase (sulfotransferase a) cataly  
A:Reference number: A37395; MUID:90294808; PMID:2113604  
A:Accession: A37395  
A:Molecule type: protein  
A:Residues: 2-21 <OGU4>  
A:Cross-references: UNIPARC:UPI0000175821  
R:Takahashi, M.; Homma, H.; Matsui, M.  
Biochem. J. 293, 795-800, 1993  
A:Title: Developmental changes in the isoelectric variants of rat hepatic hydroxysteroid  
A:Reference number: S35600; MUID:93356746; PMID:8352748  
A:Accession: S35600  
A:Molecule type: protein  
A:Residues: 2-43 <TAK1>  
A:Cross-references: UNIPARC:UPI0000175822  
A:Experimental source: PI 7.2 variant  
A:Accession: S35601  
A:Molecule type: protein  
A:Residues: 2-46 <TAK2>  
A:Cross-references: UNIPARC:UPI0000175823  
A:Experimental source: PI 6.7 variant  
R:Barnes, S.; Buchina, E.S.; King, R.J.; McBurnett, T.; Taylor, K.B.  
J. Lipid Res. 30, 529-540, 1989  
A:Title: Bile acid sulfotransferase I from rat liver sulfates bile acids and 3-hydroxy s

A:Reference number: A60949; MUID:89328199; PMID:2754334  
A:Accession: A60949  
A:Molecule type: protein  
A:Residues: 2-17, 'S', 19-23, 'D', 25, 'NKS', 31-32, 'DX', 35-37 <BAR>  
A:Cross-references: UNIPARC:UPI0000175824  
R:Song, C.S.; Kim, J.M.; Roy, A.K.; Chatterjee, B.  
Biochemistry 29, 542-551, 1990  
A:Title: Structure and regulation of the senescence marker 2 gene promoter gene promoter  
A:Reference number: I52408; MUID:90148982; PMID:2302387  
A:Accession: I52408  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-67 <SON>  
A:Cross-references: UNIPARC:UPI00000E763C; GB:M29301; NID:g206993; PIDN:AAA42151.1; PID  
C:Comment: This sulfotransferase plays roles in steroid metabolism and in the modificati  
nzyme.  
C:Genetics:  
A:Introns: 45/1  
C:Keywords: carcinogenesis; sulfotransferase  
Query Match 23.5%; Score 359.5; DB 2; Length 284;  
Best Local Similarity 33.5%; Pred. No. 1.8e-23;  
Matches 92; Conservative 53; Mismatches 105; Indels 25; Gaps 8;  
QY 19 YEFPHGVLPBFCRCK--MEIAN-PPVPSDVWIVTYPKSCTSLLOEVVYLVSOADPD 75  
Db 4 YTWEGIPFPAGPIKETLQNVKVFVKEEDLILLYPKSGTGNWLIRIVCLIQTKGDPK 63  
QY 76 EIGLMNIDEQLPVLEYPOQGLDIIKELTSPRLIKSHLPVRFPLPSDLHNGDSKVIYMARNP 135  
Db 64 WIQSVTIWDRFWIE-TDLGYDMLIKKGPRLITSHLPMLHFSKSLFSSKAKVIYLRNP 122  
QY 136 KDLVSVYYOFRSLRTMSYR---GTQFQFCRRFMNDKLGYSWFEHQVFEHWRMDSNV 191  
Db 123 RDLVSVGYFW-GKTTLAKKPSLGTVEW---FLKGNVPYGSWFEHQVFEHWRMDSNV 178  
QY 192 LFLKYEDMHRDLVTWVEQLARFVSCDKAQLALTEHCHQLVDQCCNAE-----A 242  
Db 179 LLLVYEDMKDWTGSIKKICDFLCKLEPDELNLVLKYSSFQVVKNMNSYSLMEKEI 238  
QY 243 LP---VGRGRVGLWKDIFTVSNNEKFDLVYKQKM 273  
Db 239 LPGFTFMNGTIGDKNHFHTVAQAEAFDKVQEKM 273  
RESULT 35  
T10086  
alcohol sulfotransferase (EC 2.8.2.2) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 16-Jul-1999 #sequence revision 16-Jul-1999 #text\_change 09-Jul-2004  
C:Accession: T10086  
R:Kong, A.N.T.  
submitted to the EMBL Data Library, June 1995  
A:Description: CDNA cloning of the mouse liver alcohol/hydroxysteroid form of (mSta2) su  
A:Reference number: Z16940  
A:Accession: T10086  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-285 <KON>  
A:Cross-references: UNIPROT:P50236; UNIPARC:UPI0000027B8E; EMBL:L27121; NID:g496151; PI  
A:Experimental source: strain BALB/c, liver  
C:Genetics:  
A:Gene: Sta2  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: sulfotransferase  
Query Match 23.3%; Score 356.5; DB 2; Length 285;  
Best Local Similarity 32.8%; Pred. No. 3.2e-23;  
Matches 90; Conservative 51; Mismatches 114; Indels 19; Gaps 6;  
QY 17 SKYFEFHGVRLP--PFCRGMEEIAN-PPVPSDVWIVTYPKSCTSLLOEVVYLVSOQAD 73  
Db 3 SDYNWFEIGPPIPAISYREILIRNKVFVKEEDLILTYPKSGTGNWLIRIVCLIQTKGD 62



QY 74 PDEIGLNMIDQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMAR 133  
DB 63 PKMIQTPWDRSPWIE-TEIGYSAINKEGRGLITSHLPFLHFSKSFSSKAKAIYLMR 121  
QY 134 NPKDLVSVYQYHRSRLTMSVGTGTFQFCRRPMNDKLGYSWFHVBQBFWEHRMDSNVLF 193  
DB 122 NPRDLVSGYFWGNTLVNKPNGSLGTYPFNLQNVLFSGWFHVRGWSLNRWDNVLV 181  
QY 194 LKYEDMHRDLVTMVQLARFLGVSCDKAQLAEALTEHC-----HQLV--DQCCN 239  
DB 182 LYYEDMKDXTGTTIKKICDFGLKNLGPDELVLVKYSSFOAMKNNMNSYSLIKEDRVTN 241  
QY 240 AEALPVGRGVLWKDIFTVSNNEKFDLVYKQKM 273  
DB 242 GLKL-MRKGTTGDWKNHFTVAQAEAFDKVFQEKM 274

RESULT 36

138548

alcohol sulfotransferase (EC 2.8.2.2) - human

N:Alternate names: dehydroepiandrosterone sulfotransferase (DHEA-ST); dehydroepiandroste

C:Species: Homo sapiens (man)

C>Date: 06-Sep-1996 #sequence revision 06-Sep-1996 #text change 09-Jul-2004

C:Accession: I53037; I38548; I38549; JCI1223; S28155; A58735; I37285; S52476

R:Otterness, D.M.; Wieben, E.D.; Wood, T.C.; Watson, R.W.G.; Madden, B.J.; McCormick, D.

Mol. Pharmacol. 41, 865-872, 1992

R:Luu-The, V.; Dufort, I.; Paquet, N.; Reimnitz, G.; Labrie, F.

DNA Cell Biol. 14, 511-518, 1995

A:Title: Structural characterization and expression of the human dehydroepiandrosterone

A:Reference number: I53037; MUID:95322029; PMID:759806

A:Accession: I53037

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-285 <LNU>

A:CROSS-references: UNIPROT:Q06520; UNIPARC:UPI000013C7E8; GB:I36196; NID:G908766; PIDN:

R:Otterness, D.M.; Wieben, E.D.; Wood, T.C.; Watson, R.W.G.; Madden, B.J.; McCormick, D.

Mol. Pharmacol. 41, 865-872, 1992

A:Title: Human liver dehydroepiandrosterone sulfotransferase: Molecular cloning and expr

A:Reference number: I38548; MUID:92269778; PMID:1586921

A:Accession: I38548

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-285 <OTT1>

A:CROSS-references: UNIPARC:UPI000013C7E8; EMBL:U08024; NID:G468250; PIDN:AAA17749.1; PI

A:Accession: I38549

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-285 <OT2>

A:CROSS-references: UNIPARC:UPI000013C7E8; EMBL:U08025; NID:G468252; PIDN:AAA17750.1; PI

A>Note: parts of this sequence were determined by protein sequencing; the amino end of b

A>Note: three electrophoretic forms were observed

R:Kong, A.N.T.; Yang, L.; Ma, M.; Rao, D.; Bjornsson, T.D.

Biochem. Biophys. Res. Commun. 187, 448-454, 1992

A:Title: Molecular cloning of the alcohol/hydroxysteroid form (hSta) of sulfotransferase

A:Reference number: JCI1223; MUID:92392364; PMID:1520333

A:Accession: JCI1223

A:Molecule type: mRNA

A:Residues: 1-158, 'V', 160-285 <KON>

A:CROSS-references: UNIPARC:UPI000016B38E; GB:S43859; NID:G255072

A:Experimental source: liver

A>Note: the authors translated the codon AGC for residue 222 as Thr

R:Coner, K.A.; Falany, J.L.; Falany, C.N.

Biochem. J. 289, 233-240, 1993

A:Title: Cloning and expression of human liver dehydroepiandrosterone sulphotransferase.

A:Reference number: S28155; MUID:93143674; PMID:7678732

A:Accession: S28155

A:Molecule type: mRNA

A:Residues: 1-89, 'S', 91-285 <COM>

A:CROSS-references: UNIPARC:UPI000016183A; GB:X70222; GB:S53620; NID:G312804; PIDN:CAA49

A>Note: parts of this sequence were determined by protein sequencing; the amino end of b

R:Forbes, K.J.; Hagen, M.; Glatt, H.; Hume, R.; Coughtrie, M.W.

Mol. Cell. Endocrinol. 112, 53-60, 1995

A:Title: Human fetal adrenal hydroxysteroid sulphotransferase: cDNA cloning, stable expr

A:Reference number: A58735; MUID:96034512; PMID:7589785

A:Accession: A58735

A:Molecule type: mRNA

A:Residues: 1-285 <FOR>

A:CROSS-references: UNIPARC:UPI000013C7E8; GB:X84816; NID:G683577; PIDN:CAA59274.1; PID:

A:Experimental source: fetal adrenal

R:Otterness, D.M.; Her, C.; Aksoy, S.; Kimura, S.; Wieben, E.D.; Weinshilboum, R.M.

DNA Cell Biol. 14, 331-341, 1995

A:Title: Human dehydroepiandrosterone sulfotransferase gene: molecular cloning and struc

A:Reference number: I37285; MUID:95225980; PMID:7710689

A:Accession: I37285

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-285 <RES>

A:CROSS-references: UNIPARC:UPI000013C7E8; EMBL:U13061; NID:G806716; PIDN:AAC51353.1; PI

C:Genetics:

A:Gene: GDB:STD; HST

A:CROSS-references: GDB:I32655; OMIM:125263

A:Map position: 19q13.3-19q13.3

A:Introns: 46/1; 115/3; 158/1; 189/3; 249/1

C:Complex: homodimer

C:Function:

A:Description: catalyzes formation of the sulfate esters of alcohols using 3'-phosphoaden

A:Pathway: steroid metabolism; detoxification

A>Note: sulfates bile acids and steroids, in particular dehydroepiandrosterone, in the l

C:Superfamily: alcohol sulfotransferase

C:Keywords: acetylated amino end; detoxification; homodimer; steroid metabolism; sulfotr

F:2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted

Query Match 23.2%; Score 355.5; DB 1; Length 285;

Best Local Similarity 31.0%; Pred. No. 4e-23;

Matches 83; Conservative 59; Mismatches 109; Indels 17; Gaps 5;

QY 22 FHGVRLPP-PCRGMKEIAN-FVPRPSDWIVTYPKSGTSLLOEVVVLVSGQADPDBIG 78

DB 8 PEGIAFTPMGRSETLRVDFEVRDEVDVILTYPKSGTNLWLAELCLMHSKGDAKWIIQ 67

QY 79 LMNIDQQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDL 138

DB 68 SVPIWERSPWVE-SEIGYATLSETSPRLFSHLPQLFPKSPFSSKAKVIYLMRNPDRV 126

QY 139 VVSYYQFHRSRLTMSVGTGTFQFCRRPMNDKLGYSWFHVBQBFWEHRMDSNVLFKYED 198

DB 127 LVSGYFFWKNMKFIKPKKSWEYFEWFCQGTVLGYSWFDHIGWMPMREKNFLLLSYEE 186

QY 199 MHRDLVTMVQLARFLGVSCDKAQLAEALTEHC---HQLVDOCCNAEALPVG----- 246

DB 187 LKQDTGRTIEKICDFGLKTLPEELNLILKNSFSQSMKENKMSYSLLSVDYVVDKAQL 246

QY 247 -RGRVGLWKDIFTVSMNEKFDLVYKQKM 273

DB 247 RKGSGDWKNHFTVAQAEAFDKLFQEKM 274

RESULT 37

I52849

alcohol sulfotransferase (EC 2.8.2.2) - rat

N:Alternate names: hydroxysteroid sulfotransferase

C:Species: Rattus norvegicus (Norway rat)

C>Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 09-Jul-2004

C:Accession: I52849; I65759

R:Watabe, T.; Ogura, K.; Sateukawa, M.; Okuda, H.; Hiratsuka, A.

Chem. Biol. Interact. 92, 87-105, 1994

A:Title: Molecular cloning and functions of rat liver hydroxysteroid sulfotransferases

A:Reference number: I52849; MUID:94306585; PMID:8033273

A:Accession: I52849

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-284 <RES>

A:CROSS-references: UNIPROT:P15709; UNIPARC:UPI0000170BD9; GB:D14987; NID:G303797; PIDN:

C:Superfamily: alcohol sulfotransferase

C:Keywords: sulfotransferase

Query Match 23.1%; Score 353.5; DB 2; Length 284;



```

Best Local Similarity 32.5%, Pred. No. 5.9e-23;
Matches 89; Conservative 54; Mismatches 108; Indels 23; Gaps 7;

Qy 19 YFEFHGVLRLPPFCRGK--MBEIAN-FPVRPSDVMIVYPKSGTSLQEVVVLVSQCADPD 75
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 YTWPEGIPFHAFGISKETLQVNCNKFPVKEEDLILAYPKSGTNWLLIEIVCLIQKGDPK 63
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 76 EIGLMTDEQLPVLEYQPGLDIIKELTSPLIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 64 WIQSVTIWDRSPWIE--TDVGYDILIKKGPRLMTSHLPMLFSLFSSKAKVIYLIRNP 122
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 136 KDLVVSYYQFHRs---LRTWSYRGTFOEFCRRFMNDKLGYGSMFHFVQEFWEHRMDSNVL 192
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 RDVLVSGYTFWGNSTLVKKPDSLGTYYEW---FLKGNVLYGSMFHFIRAWLSWREWDNFL 179
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 193 FLKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCHQLV---DQCCNAEALP--- 244
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 LLVYEDMKDKTMGTGIKKICDFLGKLEPDELVLVKYSSFQVMKENDMSNYSLLMKKSIF 239
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 245 -----VCRGAVGLWKDIFTYSMNEKFDLVYKQM 273
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 TGIGLMRKGTIGDWKNHFTVSAEAFDKVFQEXM 273
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 38
A33569
alcohol sulfotransferase (EC 2.8.2.2) - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 09-Jul-2004
C/Accession: A33569
R/Ogura, K.; Kajita, J.; Narihata, H.; Watabe, T.; Ozawa, S.; Nagata, K.; Yamazoe, Y.; K.
Biochem. Biophys. Res. Commun. 165, 168-174, 1989
A/Title: Cloning and sequence analysis of a rat liver cDNA encoding hydroxysteroid sulf-
A/Reference number: A33569; MUID: 90073705; PMID:2590219
A/Accession: A33569
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-284 <OGU>
A/Cross-references: UNIPROT:P15709; UNIPARC:UP1000001709B6; GB:M31363; NID:G204670; PIDN:
C/Superfamily: alcohol sulfotransferase
C/Keywords: sulfotransferase

```

Query Match	23.0%;	Score 351.5;	DB 2;	Length 284;
Best Local Similarity	32.7%;	Pred. NO. 8.7e-23;		
Matches	90;	Conservative 55;	Mismatches 105;	Indels 25; Gaps 8;
Qy	19	YFEFHGVRLPPFCGCK--MESIAN-FPVRPSPDMIVTPYKSGTSLLOEWWVLVSGADPD	75	
Db	4	YTFEGPIPFHAFGISKETQLQNVCKNFVVKDELLILAYPKSGTGNWLIIEIVCLIQTKGDPK	63	
Qy	76	EIGLWNTDEQLPVLVEYPOPGDDIINKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNP	135	
Db	64	WIQSVTTWDRSPWIE--TDVGVDIILIKKGPRKMTSHLPWHLFSKLSFSKAKVILYLRNP	122	
Qy	136	KDLVVSYVQFHRSLRTMSYR---GTFOEFCRRFMNDKLGYGSWFEHFQEFWEHRMDSNV	191	
Db	123	RDVLVSGVYFWGN--STLAKKSDSLGTYVEW---FLKGNVLXGSWFEHIRAWLSMQEWNF	178	
Qy	192	LFLKYEDMHRDLVTWVEQLARFLGVSCDKAOLEALTECHOLV---DOCCNAEALP----	244	
Db	179	LLLYEEDMKDGTMTGIKKICDFLGKLPDELDLVLYKSSFQVMKENDMSVSLIMKKSI	238	
Qy	245	-----VGRGRVGLWKDLFTVSMNEKFDLVYKQM	273	
Db	239	FTGTGLMRKGTVGDKNKHFTVSOAEAFDKVFQEKM	273	

```

RESULT 39
A26136
      senescence marker protein 2, hepatic - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 09-Jul-2004
C/Accession: A26136

```

R:Chatterjee, B.; Majumdar, D.; Ozbilen, O.; Murty, C.V.R.; Roy, A.K.  
J. Biol. Chem. 262, 822-825, 1987  
A:Title: Molecular cloning and characterization of cDNA for androgen-repressible rat liver alcohol sulfotransferase  
A:Reference number: A26136; MUID:87109186; PMID:3805009  
A:Accession: A26136  
A:Molecule type: mRNA  
A:Residues: 1-282 <CHA>  
A:Cross-references: UNIPROT:P07631; UNIPARC:UPI000017581B  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: liver

Query Match	22.18;	Score 338;	DB 2;	Length 282;
Best Local Similarity	32.08;	Pred. No. 1.3e-21;		
Matches 89; Conservative	47;	Mismatches 112;	Indels 30;	Gaps 6
Qy	17	SKYFEHGVRLP--PFCGRGMEIAN-FPYRPSDWITVPKSGTSLQVEVYLVSQAG	73	
Db	3	SDYNWPGIPPAISYQREILEDIRNFVVKEDLLILTYPKSGTNMLNEIVLCIIQTGKD	62	
Qy	74	PDEI-----GLMNIDEQLPVLEYQPQGLDIIKELTSRLTKSHLPYRFLPSDLHNGSKV	128	
Db	63	PKWIQSCPGTGVTPDELEWIFRNHGG-----PRLITSHUPIHLFSKSPFFSSKAKA	113	
Qy	129	IYNARNPKDLVSYQYFHRSLRTMSYRGTFQFCFRRFMNDKLGYSGFHFHQVEFWEHRMD	188	
Db	114	IYLMRNPRDILVSGYFFWGNNTLVKNPGLSGITYPEWFLQGNVLFGSNFVHVRGWLMSREW	173	
Qy	189	SNVLFLKYEDMHRDLVTMVQLARFLGVSCDKAQLEALTEHC-----HQLVDQ	236	
Db	174	DNFLVLYEDMKDGTMTGIIKKICDPLGNLGPDELDLLLSYFSQAMKENNMSYSLIKE	233	
Qy	237	CCNAEALPVGR-GRVGLWKDIFTVSMNEKEFDLVYKQRM	273	
Db	234	DPILITGLKLMRKGTTGDKWNHFTVAOAEAFDKYFOEKM	271	

RESULT 40

B40216

Flavonol 3'-sulfotransferase - *Flaveria chloraefolia*

C/Dates: *Flaveria chloraefolia*

C/Spec: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C/Accession: B40216

R/Varin, L.; Deluca, V.; Ibrahim, R.K.; Brisson, N.

Proc. Natl. Acad. Sci. U.S.A. 89, 1286-1290, 1992

A/Title: Molecular characterization of two plant flavonol sulfotransferases

A/Reference number: A40216; PMID:1741382

A/Accession: B40216

A/Status: preliminary; not compared with conceptual translation

A/Molecule type: nucleic acid

A/Residues: 1-311 aa>

A/Cross-references: UNIPROT:P52836; UNIPARC:UPI000012A3FC; GB:M841135; NID:19

A/Experimental source: terminal bud

A/Note: sequence extracted from NCBI backbone (NCBIP:82218)

C/Superfamily: alcohol sulfotransferase

Query Match	19.8%	Score 303	DB 2	Length 311
Best Local Similarity	28.3%	Pred. No. 1.5e-18		
Matches	82	Conservative	57	Mismatches 95; Indels 56; Gaps 11
Qy	30	FCRKMEEIANFVRPSDVMIVTPSGTSGTLLQEVVYLVSGQADPDIG-LMNI--DEQL	86	
Db	34	FLEGRILSEQFAHPNDVFLASYPGSGTTLWAKWICITIREKFPDDSTSLTLTTPHDCI	93	
Qy	87	PVLEYQPGDIIKE-----LTPSLIKSHLPYRFLPSDLHNGDSKVYMARPNKDLVWS	141	
Db	94	PLLE---KDLKLEQENQNSLYTP--ISTHFYKSPESARTSNCKIVYIRNMKVIVS	148	
Qy	142	YYQFHSRLPTMS-----YRGTQFQPCRFRMNDKLGYSWFEEHQEFWEHRMDSN--VLFL	194	
Db	149	YYHFLRQIVKLSVEEAPFEEAFDFECQIGSS---CGFYWEHIKGYWKASLEKPEIFLFL	204	
Qy	195	KYEDMHRDLVTYVEQARFLG-----VSCDRAQKLEAL-----TE	228	

Db 205 KYEDMKDPVSVKKLADFIHPFTPEKEEAGVIEDIVKLSFEKLSLEVNSKGMHRPE 264

Qy 229 HCHQLVDQCCNAEALPVGGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDL 278

Db 265 EAHSIENR-----LYFRKRGDGMKWNFTTDMTQKIDKLIDKLGATGL 308

RESULT 41

T47448

Query Match 19.7%; Score 302; DB 2; Length 329;

Best Local Similarity 26.8%; Pred. No. 2e-18;

Matches 89; Conservative 61; Mismatches 108; Indels 74; Gaps 15;

Qy 3 ESEATPTGPEFESKYFEFGVRL-----PPFCRGKMEETANFPVRPSDVWIVTY 53

Db 18 ESKTLISLPSDKNST-----GVNCKYQGCWYTPPIQLQGVNLFQKFPQDQTDIIASF 72

Qy 54 PKSGTSLQEVVY-LVSGADPDEIGLNMIDQLPVL-EYP---OPGL-----D 97

Db 73 PKCGTTLKALTFAIVRSKIPSH-----DDHFLSDGNPHVLSPEMLVYLCSNPD 126

Qy 98 IIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVSVYQFHRSLRT-----M 152

Db 127 LTKFSSSSRLFSHTMPSHTLQBLGKSTCKIVYMSRNKVDLVSVYHFFCKQKQTDNIIIS 186

Qy 153 SYRGTFQFCR--RPMNDKLGVSWEFHVQFWEHRMD--SNVFLKYEDMHRDLVTWVE 208

Db 187 SVEDTFEMFCRGVNF-----FGPFMDHVLVSYWRGSLDPNHNLFMPKPEEMKEPREQIK 240

Qy 209 QLAREFLGVSCDKAOLEA-LTEHCHQLVDQCC--NAEALPVG-----RG 248

Db 241 RLAEFLGCLFTKEEESGLVD---EIIDLSLRNLSLEINKTKGLHSTGRENKTFPRKG 297

Qy 249 RVGLMKDIFTVSMNEKFDLVYKQKMGKCDLTF 280

Db 298 EVGDMKNVLTPEMNKIDMIQELQNSGLKF 329

RESULT 42

B84452

Query Match 18.6%; Score 284.5; DB 2; Length 323;

Best Local Similarity 25.6%; Pred. No. 6.4e-17;

Matches 74; Conservative 58; Mismatches 90; Indels 67; Gaps 11;

Qy 40 NFPVRPSDVWIVTYPKSGTSLQEVVY-LVSGADPDEIGLNMIDQLPVL----- 89

Db 54 SFKPDQTDIIIVASPKCGTTLKALTFAIVRSKIPSH-----DDHFLSDGNPHVLP 107

Qy 90 -----BYPOPGDLIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLV 139

Db 108 YFEIDLVLRSNP-----DLTKFSSSPRLFSHTVPSHTLQELGKSTCKIVYISRNVDTL 163

Qy 140 VSYVQFHRSLT-----MSYRGTFQFCRPFMDKLG---YGSNFEHVQFWEHRMD--S 189

Db 164 VSYWHFFTKQTDKLISSFEDTFEMFCR-----GVSIFGPFMDHVLVSYWRGSLDPN 216

Qy 190 NVLFLKYEDMHRDLVTWVEQLARFLGVSCDKAOLEALTEHCHQLVDQCC--NAEALPVG- 246

A;Cross-references: UNIPROT:Q9ZPQ5; UNIPARC:UPI00000A21B0; GB:AE002093; NID:G4406768; PI

C;Genetics:

A;Gene: At3g03770

A;Map position: 2

C;Superfamily: alcohol sulfotransferase

Query Match 19.4%; Score 297.5; DB 2; Length 324;

Best Local Similarity 25.8%; Pred. No. 4.8e-18;

Matches 84; Conservative 57; Mismatches 124; Indels 61; Gaps 10;

Qy 1 MAESE-----AETPTGPEFESKYFEFGVRLPP-FCRGKMEETANFPVRPSDVWI 50

Db 12 LVESELVQCEBELLSSLPDRSRVFAEYLQYQGFYPPENLLEGLVYSQKHFQARDSDIVL 71

Qy 51 VTYPKSGTSLQEVVYLVSGADPDEIGLNMIDE-QLPVLEYPO-----PGLD 97

Db 72 ASIPKSGTTLWLSLVF-----ALIHQBFQTPLVSHPLLDNNPHVTLVTFIEGFH 120

Qy 98 IIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVSVYQFHRSLRTMSVRG- 156

Db 121 LHTQDTPRIESTHIPVGSLPESVKDSCKVYCCRNPKDAFVSLWHFMKNLIVKEMVGC 180

Qy 157 TFQFPCRRFMNDKLGVSWEFHVQFWEHRMDS--NVFLKYEDMHRDLVTWVEQLARFL 214

Db 181 TMEEMVRFFCRGSSIVGPFMDHVLQYWKESRENPKVMFVMEYEMREQOPQEWVMRIAEFL 240

Qy 215 GVSCDKAOLE-ALTEHCHQLVDQCCNAEALP-----VGRGRVGLWK 254

Db 241 GYSFTEEEIENGVLIEDIKL---CSLENLSKLEVNEKGKLLNGMETKAFPRKGEIGGWR 296

Qy 255 DIFTVSMNEKFDLVYKQKMGKCDLTF 280

Db 297 DTLPLLAEEIDTKTEKLGSDPRF 322

RESULT 43

T47447

Query Match 18.6%; Score 284.5; DB 2; Length 323;

Best Local Similarity 25.6%; Pred. No. 6.4e-17;

Matches 74; Conservative 58; Mismatches 90; Indels 67; Gaps 11;

Qy 40 NFPVRPSDVWIVTYPKSGTSLQEVVY-LVSGADPDEIGLNMIDQLPVL----- 89

Db 54 SFKPDQTDIIIVASPKCGTTLKALTFAIVRSKIPSH-----DDHFLSDGNPHVLP 107

Qy 90 -----BYPOPGDLIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLV 139

Db 108 YFEIDLVLRSNP-----DLTKFSSSPRLFSHTVPSHTLQELGKSTCKIVYISRNVDTL 163

Qy 140 VSYVQFHRSLT-----MSYRGTFQFCRPFMDKLG---YGSNFEHVQFWEHRMD--S 189

Db 164 VSYWHFFTKQTDKLISSFEDTFEMFCR-----GVSIFGPFMDHVLVSYWRGSLDPN 216

Qy 190 NVLFLKYEDMHRDLVTWVEQLARFLGVSCDKAOLEALTEHCHQLVDQCC--NAEALPVG- 246

A;Cross-references: UNIPROT:Q9MIV2; UNIPARC:UPI00000A990C; EMBL:AL138649

C;Genetics:

A;Map position: 3

A;Note: T14D3.10

C;Superfamily: alcohol sulfotransferase

Query Match 18.6%; Score 284.5; DB 2; Length 323;

Best Local Similarity 25.6%; Pred. No. 6.4e-17;

Matches 74; Conservative 58; Mismatches 90; Indels 67; Gaps 11;

Qy 40 NFPVRPSDVWIVTYPKSGTSLQEVVY-LVSGADPDEIGLNMIDQLPVL----- 89

Db 54 SFKPDQTDIIIVASPKCGTTLKALTFAIVRSKIPSH-----DDHFLSDGNPHVLP 107

Qy 90 -----BYPOPGDLIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLV 139

Db 108 YFEIDLVLRSNP-----DLTKFSSSPRLFSHTVPSHTLQELGKSTCKIVYISRNVDTL 163

Qy 140 VSYVQFHRSLT-----MSYRGTFQFCRPFMDKLG---YGSNFEHVQFWEHRMD--S 189

Db 164 VSYWHFFTKQTDKLISSFEDTFEMFCR-----GVSIFGPFMDHVLVSYWRGSLDPN 216

Qy 190 NVLFLKYEDMHRDLVTWVEQLARFLGVSCDKAOLEALTEHCHQLVDQCC--NAEALPVG- 246

Db 217 HVLFMKFEEMKAPRDQIKKFAEFLG--CPFTKEEESGSVDIEIIDLCSRLNLSLEINK 274  
Qy 247 -----RGRVGLWKDIFTVSMNEKFDLVVYKQKMGKCDLTF 280  
Db 275 TGKLNSGRENKMFPRKGEVGDWKNVLTPEMKNKIDWIOEKLQNSGLKP 323

RESULT 44  
T07833  
probable steroid sulfotransferase (EC 2.8.2.15) 3 - rape  
C:Species: Brassica napus (rape)  
C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 09-Jul-2004  
C:Accession: T07833  
R:Richard, M.; Nicolle, L.; Varin, L.  
submitted to the EMBL Data Library, April 1997  
A:Reference number: Z16161  
A:Accession: T07833  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-325 <RIC>  
A:Cross-references: UNIPROT:O82410; UNIPARC:UPI00000AA137; EMBL:AF000307; NID:G3420007;  
C:Genetics:  
A:Gene: ST3  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: sulfotransferase

Query Match 18.3%; Score 280; DB 2; Length 325;  
Best Local Similarity 26.6%; Pred. No. 1.6e-16;  
Matches 82; Conservative 60; Mismatches 118; Indels 48; Gaps 12;

Qy 9 PSTGPEFSKYFEFHG-VRLPFCRGKMBEIANFFVRPSDWMIVTPKSGTSLLOEVVYL 67  
Db 28 PSEKGLVSIQYQFQGRWHTALLQGITCQKHFAKSDIILVTNPKSGTTWLKSLVFA 87  
Qy 68 -----VSQADPDEIGLMIIDQLPVLE---YPOGLDIIKELTSPRLIKSHLPFRF 116  
Db 88 LINRHKFPVSSGDHP--LLVTNPHLLVPMFMEGVYESPDF--FSLPLPRLMNTSHLS 144

Qy 117 LPSDLHGDSKVIYMARPNKDLVSYQFHRSL---RTMSYRGTFQFCRRRPMNDKLGVG 173  
Db 145 LPESVSKSSCOIYVCNPKDMFVSLWHFGKKLAPQETADY--PLEKAVEAFCQCKFTAG 202

Qy 174 SWFHEVQFWEHRMD--SNVLFLKYEDMHRDLVTWQELARFLGVSCDKAQLEALTEHCH 231  
Db 203 PFWDHLEVWYASLEPNKVLFTVTEELKKQFEVEKRIAEFIG--CGTAEVEVSE--- 257

Qy 232 QLVDQCCNAEA-----LPVG-----RGRVLWKDIFTVSMNEKFDLVYKQK 272  
Db 258 --IVKLCSPESLSRLEVNROGKLPNGIETNAPFRKGEIGGWRDTLSESLAIDRTTBK 315

Qy 273 MGKCDLTF 280  
Db 316 FGGSGLKPF 323

RESULT 45  
E86319  
probable flavonol sulfotransferase [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: E86319  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: AB6141; MUID:21016719; PMID:11130712

A:Accession: E86319  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-346 <STO>  
A:Cross-references: UNIPROT:Q9FZ80; UNIPARC:UPI000000C631; GB:AE005172; NID:G9795597;  
C:Genetics:  
A:Map position: 1  
C:Superfamily: alcohol sulfotransferase

Query Match 18.2%; Score 278.5; DB 2; Length 346;  
Best Local Similarity 24.9%; Pred. No. 2.3e-16;  
Matches 83; Conservative 62; Mismatches 119; Indels 69; Gaps 12;

Qy 1 MAESEAE-TPSTPGEFE--SKYFEPHGVRLP-----PFCRGKME 36  
Db 11 VSESHNELASSSPSEFEKNQKHQEIATLPHKDGWRPKDPFVEYGGHWWLQPLEGLLH 70

Qy 37 EIANFPVRPSDWMIVTPKSGTSLLOEVVYLSQCADPD-----EIGLMIIDQLPVLE-- 90  
Db 71 AQKFKARPNDFFVCSYPKTGTTLKALTFAIANRKSFDVSTNPLKRNPHFEFVYIEID 130

Qy 91 YP-QPGLDIIKELTSPRLIKSHLPYRFLPSDLHGDSKVIYMARPNKDLVSYQFHRSL 149  
Db 131 PFPFSDVVLKD-EGNTLFTSTHPIYDILLPESVVGCKIVYWRDPKDTFVSMWTFAPKE 189

Qy 150 RT-----MSYRGTFQFCRRRPMNDKLGYSWFHVEHQEFWE--HRMDSNVLFYKEDMHRD 202  
Db 190 RSQGPVVSIEEAFDKYCQGLS----AYGPYLDHVLGYKAYQANPDQILFLKYETMRAD 245

Qy 203 LVTWQELARFLGVSCDKAQLEALTEHCHQLVDQCCNAEALP----- 244  
Db 246 PLPYVKRLAEAFMGYGTKEEBEG---NVVEKVKVLCSPETLKNLEANKGEKDRDPVY 302

Qy 245 -----VGRGRVLWKDIFTVSMNEKFDLVYKQK 272  
Db 303 ANSAVFRKGVQDWQNYLTPEWVARIDGLMEEK 335

RESULT 46  
AB4452  
probable steroid sulfotransferase [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: AB4452  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umavam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: AB4420; MUID:20083487; PMID:10617197  
A:Accession: AB4452  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-326 <STO>  
A:Cross-references: UNIPROT:P52839; UNIPARC:UPI00000484E4; GB:AE002093; NID:G4406767;  
C:Genetics:  
A:Gene: At2G03760  
A:Map position: 2  
C:Superfamily: alcohol sulfotransferase

Query Match 18.1%; Score 277.5; DB 2; Length 326;  
Best Local Similarity 26.5%; Pred. No. 2.6e-16;  
Matches 82; Conservative 57; Mismatches 121; Indels 49; Gaps 12;

Qy 9 PSTGPEFSKYFEFHG-VRLPFCRGKMBEIANFFVRPSDWMIVTPKSGTSLLOEVVYL 67  
Db 28 PKKXGLVSEIYFQGLMHTQAILQGLICQKFEAKDSIILVTNPKSGTTWLKALVFA 87

Qy 68 V-----SQGADPDEIGLMIIDQLPVLE---YPOGLDIIKELTSPRLIKSHLPYR 115  
Db 88 LLNRHKFPVSSGNHP--LLVTNPHLLVPMFMEGVYESPDF--FSSLSPLMNTSHLS 144

Qy 116 FLPSDLHGDSKVIYMARPNKDLVSYQFHRSL---RTMSYRGTFQFCRRRPMNDKLG 172

Db 145 SLPSVSSSKIVVCCRNPKDMFVSLWHFGKLAPEADY--PIEKAVAFCEGKFIG 202  
Qy 173 GSWFEHVQEFW--EHRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHC 230  
Db 203 GPFWDHILEYVWASRENPNKLVFTYBELKKQTEVEMKRIAFL--ECGFIEEBEVRE-- 258  
Qy 231 HOLVDQCCNAEA-----LPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQ 271  
Db 259 ---IVKLCSPFSLNLEVNKEGLPNGIETKTFPRKGBIGGWRDRTLSSSLAEIDRTIEE 315  
Qy 272 KMGKCDLTF 280  
Db 316 KFGSGGLKF 324

## RESULT 47

T07831

probable steroid sulfotransferase (EC 2.8.2.15) 1 - rape

C:Species: Brassica napus (rape)

C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 09-Jul-2004

C:Accession: T07831

R:Richard, M.; Nicolle, L.; Varin, L.

submitted to the EMBL Data Library, April 1997

A:Reference number: Z16161

A:Accession: T07831

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-323 &lt;RIC&gt;

A:Cross-references: UNIPROT:O82408; UNIPARC:UPI00000A1423; EMBL:AF000305; NID:g3420003;

C:Genetics:

A:Gene: ST1

C:Superfamily: alcohol sulfotransferase

C:Keywords: sulfotransferase

Query Match 18.1%; Score 276.5; DB 2; Length 323;  
Best Local Similarity 27.4%; Pred. No. 3.1e-16;  
Matches 85; Conservative 59; Mismatches 113; Indels 53; Gaps 14;  
Qy 9 PSTGPFESKYPEFHGV-RLPPFCRGKMEETANFFVRPSDWMIVTPYKSGTSLLOEVVYL 67  
Db 27 PSEKWLVSQMYQFEGIQTOALVQGVNVCQKHFANDSDVILATLAKSGTTWLKALLFA 86  
Qy 68 V-----SQGADPEIGIMNIDEQLPVE---YFQPGLDIIKELTSPRLIKSHLPYRFL 117  
Db 87 LIHRHKFPVSGKHP--LLVTNPHSLVPYLEGDCYSSPEVN-PAELPSRLMQTLTHHSL 143  
Qy 118 PSDLHNGDSKVIYMARNPKDLVSVYQFHRSL---RTMSY--RGTFQFCR-RFMNDKLG 171  
Db 144 PVSISKSSCKIYCCRNPKDMFVSIWHFGRKLAPEKTAEPYPIETAVAAFCCKKFIG--- 199  
Qy 172 YGSMFEHVQEFWEHRM--DSNVLFLLKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHC 229  
Db 200 -GPFWDHILEYVWASRENPNKLVFTYBELKKQTEVEMKRIAFL--ECGFIEEBEVRE- 255  
Qy 230 CHOLVDQCCNAEA-----LPVG-----RGRVGLWKDIFTVSMNEKFDLVYK 270  
Db 256 ---IVKLCSPFSLNLEVNROGKLPNGIESNAFFRKGTGGRDRTLSSSLADVIDRTTE 311  
Qy 271 QMGKCDLTF 280  
Db 312 QKFGSGGLKF 321

## RESULT 48

A40216

flavonol 4'-sulfotransferase - Flaveria chloraefolia

C:Species: Flaveria chloraefolia

C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Sep-1999

C:Accession: A40216

R:Varin, L.; Deluca, V.; Ibrahim, R.K.; Brisson, N.

Proc. Natl. Acad. Sci. U.S.A. 89, 1286-1290, 1992

A&gt;Title: Molecular characterization of two plant flavonol sulfotransferases.

A:Reference number: A40216; MUID:92159034; PMID:1741382

A:Accession: A40216

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-320 &lt;VAR&gt;

A:Cross-references: UNIPARC:UPI0000175837

A:Experimental source: terminal bud

A:Note: sequence inconsistent with the nucleotide translation

A:Note: sequence extracted from NCBI backbone (NCBIN:82216, NCBI:82217)

C:Superfamily: alcohol sulfotransferase

Query Match 18.0%; Score 275.5; DB 2; Length 320;  
Best Local Similarity 27.5%; Pred. No. 3.8e-16;  
Matches 80; Conservative 63; Mismatches 111; Indels 37; Gaps 11;  
Qy 18 KYFEHGVRLPPFCRGKMEETANFFVRPSDWMIVTPYKSGTSLLOEVVYVLSQGADPEI 77  
Db 34 KYQDFWG--LQNNIEGAILAQOSFKARPDVFLCSYPSKSGTTWLKALAYAVTREKPDFE 91  
Qy 78 G---LMNIDEQ-LPVLEYPQGLDIIKELTSPR-----LIKSHLPYRPLPSDLHNGDSKV 128  
Db 92 TSPLLTNIPHNCPYIE-----KOLKKTVDNQNNSCFTPMATHMPYHVLPLKSLIALNCKM 146  
Qy 129 IYMARNPKDLVSVYQFHRSLRTMSY--GTQEFRCRFRMNDKLGYSGFHVFQWQFWEHRM 187  
Db 147 VYIYRNKDVIVSYFHGREITKPLEDAPFEAEFDEYHGISQGPYWDHLLGVYKASL 206  
Qy 188 D--SNVLFLLKYEDMHRDLVTMVEQLARFLGVSC-----DKAQLBALTEHC-----HOLV 234  
Db 207 ERPEVILFLKYEDVKKDPTSNVKRLAEFIGYPTFEEKEGVIESIIKLCSPENLSNLEV 266  
Qy 235 DQCCNAEA-LPV-----GGRGVGLWKDIFTVSMNEKFDLVYKQMGKCDL 278  
Db 267 NKGNSGKFLPIENLYFRKAKDGDWKNYFTDEMTKIDKIDKLSATGL 317

## RESULT 49

H96768

protein flavonol sulfotransferase F2P9.4 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: H96768

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A&gt;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: H96768

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-350 &lt;STO&gt;

A:Cross-references: UNIPROT:Q9C9C9; UNIPARC:UPI000000C528; GB:AE005173; NID:g7109463; PI

C:Genetics:

A:Gene: F2P9.4

A:Map position: 1

C:Superfamily: alcohol sulfotransferase

Query Match 17.5%; Score 267; DB 2; Length 350;  
Best Local Similarity 25.5%; Pred. No. 2.3e-15;  
Matches 83; Conservative 56; Mismatches 121; Indels 60; Gaps 12;  
Qy 2 AESEATPSTGCESEKYF-----EFHGV-RLPPFCRGKMEETAN 40  
Db 19 SHDETCTESTFEKNQRYQDLSTFPPEKGNRPKEPLIEYGGVWMLPSLEGCTHAQEF 78  
Qy 41 FVVRPSDWMIVTPYKSGTSLLOEVVYVLSQGADPE-----ICLMNIDQLPVLE--YP-Q 93

Db 79 FOARPSDFLVCSYPKTGTWTKALTAFAIANRGRFDDSSNPLLRKRNPHFVVPVIEIDFPFF 138

Qy 94 PGLDIIKELTSRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVSYQFHRSLRT-- 151

Db 139 PEVDVLKD-KGNTLFTSTHYPYELLPPDSVVKSGCKMYIWRPKDFTFISMTFLHKERTEL 197

Qy 152 ---MSYRGTFQBFRCRPMNDKLGYSWFHVFQEFWEHRMDS--NVLFLKYEDMHRDLVTM 206

Db 198 GPNVLEESFDFRCGLS----GYGPLYNHILAYWKAYQENPDRILFLKYETMRADPLPY 253

Qy 207 VEOLARFLVSC-----DKAOLEALTEHC--HQLVDCCN-----AEALPVG 246

Db 254 VKSLAEFMHGHTABEEKGVEKVVNLCSPFETLKNLEANKGEKREDRPGVYANSAYPR 313

Qy 247 RGRVGLWKDIFTVSMNEKFDLVYKQK 272

Db 314 KGKVGDSNYLTPEMAARIDGLMEEK 339

RESULT 50

T07832

probable steroid sulfotransferase (EC 2.8.2.15) 2 - rape

C:Species: Brassica napus (rape)

C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 09-Jul-2004

C:Accession: T07832

R:Richard, M.; Nicolle, L.; Varin, L.

A:Submitted to the EMBL Data Library, April 1997

A:Reference number: Z16161

A:Accession: T07832

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-324 <R1C>

A:Cross-References: UNIPROT:O82409; UNIPARC:UPI0000046BEA; EMBL:AF000306; NID:G3420005;

C:Genetics:

A:Gene: ST2

C:Superfamily: alcohol sulfotransferase

C:Keywords: sulfotransferase

Query Match 17.4%; Score 265.5; DB 2; Length 324;

Best Local Similarity 26.8%; Pred. No. 2.8e-15;

Matches 80; Conservative 61; Mismatches 128; Indels 29; Gaps 11;

Qy 9 PSTPGEFSKYFEFHG-VRLPPFCRGKMBEIANFPVRPSDVWIVYPKSGTSLLOEVY- 66

Db 28 PSEKGLVCMYQFQGRWHTQALLOGLTCQKHFEAKSDIILVTNPKSGTTWLKALVFA 87

Qy 67 LVSGADPDEIGLWNIDEQLPVLEYQPCLDIK-----ELTSPRLIKSHLPYRFLPS 119

Db 88 LINRHKFP-VYSVIIISCYQSALLVPFLGRSLRSPDFFSQSSPRLNTHISHLSLPE 146

Qy 120 DLHNGDSKVIYMARPKDLVSYQFHRSL---RTMSYRGTFQBFRCRPMNDKLGYSWF 176

Db 147 SVKSSCKIVYCRNPKDMFVSLWHFGKLAPEETADY--PIEKAVEAFCQCKFIGGPPW 204

Qy 177 EHVFQEFWEHRMD--SNVLFLKYEDMHRDLVTWVEQLARPLGVS-CDKAOLEALTEHC--H 231

Db 205 DHVLEYWYASLENPNKVLFSVSEEPKKTGETIKRIAEFLGCGLVGESEVRAIVKLCSP 264

Qy 232 QLVDQCCNAA-LPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280

Db 265 SUSSLEVNRGKLPDSGMETRAFRKGEVGGWRDITLTSUAEVIDRTIEEKFGSGGLKF 322

RESULT 51

F86407

probable sulfotransferase F3H9.17 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: F86407

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;

anen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Mafti, R.; Marziall,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: F86407

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-326 <STO>

A:Cross-References: UNIPROT:O9F291; UNIPARC:UPI00000482DF; GB:AE005172; NID:g9795621; P1

C:Genetics:

A:Map position: 1

C:Superfamily: alcohol sulfotransferase

Query Match 17.2%; Score 262.5; DB 2; Length 326;

Best Local Similarity 27.2%; Pred. No. 5.2e-15;

Matches 85; Conservative 56; Mismatches 112; Indels 59; Gaps 13;

Qy 14 EFES-KYFEFHGVRLLPPEC-----RGKMBEIANFPVRPSDVWIVYPKSGTSLLOEVV 65

Db 29 DFDSTKFKYQG-----CWYDDKTLOGVLNFORGFEPQDTIIIASFPKSGTTWLKALT 82

Qy 66 VLV-----SOGADPDEIGLWNIDEQLPVLEY-----PQCLDIKELTSRLIKSHLP 113

Db 83 VALLERSKQKHSSDDHPLLNDPHGLVFLRLFTETSKPDLTIS--SSPRLSTHVA 140

Qy 114 YRFLPSDLHNGDSKVIYMARPKDLVSYQFHRSLRTMSYR----GTQBFRCRFRMND 168

Db 141 FQTLREALNKSPCKIVYWRNVKDLVSVFWFNSAKLIEBERSILDSMFESFCRGVIN- 199

Qy 169 KLGYSWFHVFQEFWEHRMD--SNVLFLKYEDMHRDLVTWVEQLARPLGVSCKAQLAEL 226

Db 200 ---YGPSMEHVLNRYWASLEDSKNVFLKYBELKTEPRVQLKRLAEFL--DCPFTVEEBE 254

Qy 227 TEHQHQLVDQCC-----NAEALPVG-----RGRVGLWKDIFTVSMNEKFDLV 268

Db 255 RGSVEEILDCLSLRNKLNKTKTLRGADHKIFFRKGEVGDGSKNHLTPEMEKIIMI 314

Qy 269 YKQKMGKCDLTF 280

Db 315 TEEKFEGSDLK 326

RESULT 52

S69188

probable flavonol sulfotransferase (EC 2.8.2.-) - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 15-Nov-1996 #sequence\_revision 27-Feb-1997 #text\_change 21-Jul-2000

C:Accession: S69188

R:Lacomme, C.; Roby, D.

A:Title: Molecular cloning of a sulfotransferase in Arabidopsis thaliana and regulation

A:Reference number: S69188; MUID:96270377; PMID:8639757

A:Accession: S69188

A:Molecule type: mRNA

A:Residues: 1-302 <LAC>

A:Cross-References: UNIPARC:UPI0000046BEB; EMBL:Z46823; NID:g599639; PIDN:CAA86850.1; P1

C:Superfamily: alcohol sulfotransferase

C:Keywords: sulfotransferase

Query Match 17.0%; Score 259.5; DB 2; Length 302;

Best Local Similarity 26.9%; Pred. No. 8.5e-15;

Matches 77; Conservative 52; Mismatches 108; Indels 49; Gaps 12;

Qy 9 PSTPGEFSKYFEFHG-VRLPPFCRGKMBEIANFPVRPSDVWIVYPKSGTSLLOEVYL 67

Db 28 PREKGLVSEIYEFQGLWHTQAILQGLICQKRFKAKSDIILVTNPKSGTTWLKALVFA 87

Qy 68 V-----SOGADPDEIGLWNIDEQLPVLE---YPOPLDIKELTSRLIKSHLPVR 115

Db 88 LLNRHKFPVSSSGNHP--LLVTNPHLLVFLFEGVYESPDFD-FSSLSPLRLNTHSHL 144



Db 77 SGTTLWLTALFPALVQSKSLHEDHQHPILLHNPHEI-VPNLELDL-YLKSSKP--DLTKF 132

Qy 102 L-----TSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVVSYYQFHRSLR-----TM 152

Db 133 LSSSSSSPRLFSFTHMSLDLPQVLKENLCKIVVVCNVDKVMVSVWVYFQSKKIKITRAEDY 192

Qy 153 SYRGTFQBFRCRRFMDKLGYSWFEHVQFWEHRMD--SNVFLPKYEDMHRDLVTMVEOL 210

Db 193 SLEALFESFC---NGVTUHGFPWDHALSYWRGSLDDPKHFLFMRVEDLKASPTQVKRL 248

Qy 211 ARFLGVSC-----DKAQLEALTEHCHQLVDQCCNAEALPVG-----RG 248

Db 249 AEFL--DCPFTKEEDSGSVKLELC-----SLSNLRSVEINKTTSRVDKSFYRKG 301

Qy 249 RVLGWLKDIFTVSMNEKFDLVYKQMGKCDLTF 280

Db 302 QVGDWKSYPTEMVDKIDMIIEKLKGSGLKF 333

RESULT 56

T06012

hypothetical protein T25K17.90 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 09-Jul-2004

C:Accession: T06012

R:Devan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X submitted to the Protein Sequence Database, March 1999

A:Reference number: Z15382

A:Accession: T06012

A:Molecule type: DNA

A:Residues: 1-314 <REV>

A:Cross-references: UNIPROT:Q9ST06; UNIPARC:UPI00000489F1; EMBL:AL049171; GSPDB:GN00062;

A:Experimental source: cultivar Columbia; BAC clone T25K17

C:Genetics:

A:Gene: ATSP:T25K17.90

A:Map position: 4

A:Introns: 109/3

C:Superfamily: alcohol sulfotransferase

Query Match 15.4%; Score 235.5; DB 2; Length 314;

Best Local Similarity 25.6%; Pred. No. 1.1e-12;

Matches 69; Conservative 58; Mismatches 104; Indels 39; Gaps 10;

Qy 30 FCRGKMEIANTPVRPSDVWIVTPKSGTSLLOEV-VYLVSGQADPEIGLNMIDQL-- 86

Db 46 FLQGVLFNFGFGPKPQDTHIVASPKSGTLWLKALTVALFERTKNPSH-----DPMWH 99

Qy 87 PVLEYQPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVVSYYQPH 146

Db 100 PLLSNPNHNL---LSSSSPRLFSFTHPTTQVAVKDSCKVYVICRDAKSLVS--RWH 154

Qy 147 ---RSLRTMSYRGTFQFRCRRFMDKLGYSWFEHVQFWEHRMD--SNVFLPKYEDMHR 201

Db 155 IVCRSLNKEEDRTIIESMPESFCSGVCLFGPFDWHLISYWKASLEKPKQVLFMRIDEIKT 214

Qy 202 DLVTWVQELARFLGVSCDKAQLEALTEHCHQLVDQCC--NABALPVG----- 246

Db 215 DPHGQLKLAELFLG--CPSPKEEKNGSLNKILEMCSLPNLSLEVNTKGKSINGIEYKN 272

Qy 247 ---RCRGVLGWLKDIFTVSMNEKFDLVYKQMGK 273

Db 273 HPRKRGVGDWKNHLLTPEMGSKIDMIMKEKL 302

RESULT 57

H84451

probable steroid sulfotransferase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C:Accession: H84451

R:Lin, X.; Kaul, S.; Rounslev, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: H84451

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-331 <STO>

A:Cross-references: UNIPROT:Q9ZPQ6; UNIPARC:UPI00000484E3; GB:AE002093; NID:G4406766; P

C:Genetics:

A:Gene: At2G03750

A:Map position: 2

C:Superfamily: alcohol sulfotransferase

Query Match 14.7%; Score 224.5; DB 2; Length 331;

Best Local Similarity 25.1%; Pred. No. 1e-11;

Matches 68; Conservative 53; Mismatches 117; Indels 33; Gaps 8;

Qy 40 NFPVRPSDVWIVTPKSGTSLLOEVVYLVSQ-----GADPDEIGLNMIDQLPVLE--- 90

Db 63 HFKPRDTHIIILASLPKGGTTLKSLIFAVVHREKVRGTPQTTHPLLQNHDPHDLVPFLEVEL 122

Qy 91 YPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVVSYYQFHRSL- 149

Db 123 VANSQIPDLAKYSSPMIFSTHMLQAL-REATTKACTVYVCRGKIDTFVSGWHYRNMLH 181

Qy 150 RTMSYRGTFQFRCRRFMDKLGYSWFEHVQFWEHRMDS--NVLFKYEDMHRDLVTMV 207

Db 182 RTKMDQATFELMFDAYCRGVLLYGPYHEVLSYWKSGLEAKENVLFMYEETIEEPRVQV 241

Qy 208 EQLARFLGVSCDKAQLEALTEHCHQLVDQCC--NAEALPVG-----RGR 249

Db 242 KRLAEFL--ECPFTKEEESGVEEILKLCSLRNLSNLEVNKNGTTRIGVDQSVQFPRKGE 299

Qy 250 VGLMKWDIFTVSMNEKFDLVYKQMGKCDLTF 280

Db 300 VGDWKNHLLTPQMAKTFDEIIDYRLGDSGLIF 330

RESULT 58

A95936

probable alcohol sulfotransferase (EC 2.8.2.2) [imported] - Sinorhizobium meliloti (stra

C:Species: Sinorhizobium meliloti

C>Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004

C:Accession: A95936

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A:Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo

A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: A95936

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-280 <KUR>

A:Cross-references: UNIPROT:Q92VF3; UNIPARC:UPI000000CB647; GB:AL591985; PIDN:CAC49153.1;

A:Experimental source: strain 1021, megaplasmid pSymb

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.,

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: Smb21249

A:Genome: plasmid

C:Keywords: sulfotransferase

Query Match 10.5%; Score 161; DB 2; Length 280;

Best Local Similarity 23.9%; Pred. No. 2.5e-06;

Matches 64; Conservative 46; Mismatches 98; Indels 60; Gaps 12;

Qy 46 SDWIVTPKSGTSLLOEVV--YLVSGQADPEIGLNMIDQLPVLEYPQGLDIIEKLT 103





Db 140 LHEAVVPFHERIAPPPAEGLHARSPTTEP-RDWMGPNQPPGIGFTHLKGITGLANTILH 198  
Qy 181 ----EFWEHRMDSNVLFKLYEDMHRDLVTMVEQLARFLGVSCDK-----AQLEALTE 228  
Db 199 QLGTVMVRRHLENVALFHYADYQADLAGELLRPARVLGIAATRDARDLAQVATLDAMRS 258  
Qy 229 HCHQL-----VDCCNAEALPVGRGRVLWKDIFT 258  
Db 259 RASEIAPNTDGIWHSDFRFRGGSGDWQOQFFT 292

RESULT 62  
T29144  
partial CDS - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
A:Accession: T29144  
R:Pauley, A.; Gattung, S.  
submitted to the EMBL Data Library, July 1996  
A:Description: The sequence of C. elegans cosmid K11C4.  
A:Reference number: Z20577  
A:Accession: T29144  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-5107 <PAU>  
A:Cross-references: UNIPROT:Q94279; UNIPARC:UPI000011013D; EMBL:U64854; PIDN:AAB18318.1;  
A:Experimental source: strain Bristol N2; clone K11C4  
C:Genetics:  
A:Gene: CESP:unc-68  
A:Map position: 5  
A:introns: 27/1; 64/3; 92/3; 127/1; 158/2; 1222/2; 1300/2; 1347/2; 1391/1; 1419/3; 1517/3; 3269/2; 3313/2; 3466/1; 3519/3; 3629/3; 3658/2; 3710/1; 3741/3; 3779/2; 3810/3;  
C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolo

Query Match 6.6%; Score 101; DB 2; Length 5107;  
Best Local Similarity 21.7%; Pred. No. 20;  
Matches 79; Conservative 43; Mismatches 138; Indels 104; Gaps 14;

Qy 7 ETPSPGFESKYFFHGVRLPP-----FCRGK-----MEEIANFPVRPSDVIV 51  
Db 1962 QTDMPFAEAAKTKEF---RCPPKQMFRLLMCKYKEERDPELMEEDADVQCPWAEGIQ 2018  
Qy 52 TYPKSGTSLQEVVYLVSGQADPDEIGLWNIDE-----QLPVLEYQPQGLD 97  
Db 2019 QQLRDFCELLGVKIGVAKEGSDDDQLALTESEGSWVDSFARIVVKVPPVLE---EGME 2075  
Qy 98 IIKELTSP-RLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVSVYQFHSRLR-TMSVR 155  
Db 2076 MOKKGTQNFREIIVTMLREWAQDFESK-----LIRNFRLLRLQYSGIREIRAMSQT 2131  
Qy 156 GTFQFECRRFMNDKLG-----GSWFHVQE-----FWEHRMDSNVLFKLYEDMHR 201  
Db 2132 VYFHERNEKVDTFVLVIQIRELLTVQFHETEAILKRWLW--KLNNRIFFQHPDLMR 2189  
Qy 202 -----DLVTMVEQLARFLGVSC--DKAQ 222  
Db 2190 LLSVHENVMSIMMNIILTAQQTVEHEGDELKEKAPIKDAEMVAVACSRFLCYFCRTSRQN 2249  
Qy 223 LEALTEHCHQLVQCCNAEALPVGRGRVLWKDIFTVSNNKFDLVYKQK-----MG 274  
Db 2250 QKAMPEHLUSFLLDNATMLLARSPLRSGVPLDVAYSSFMDDNNELALKEELDQVAVYLS 2309  
Qy 275 KCDDL 278  
Db 2310 RCGL 2313

RESULT 63  
T03851  
thyroid hormone receptor activator molecule - human  
C:Species: Homo sapiens (man)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C:Accession: T03851; T03749; T03443  
R:Takeishi, A.; Cardona, G.R.; Koibuchi, N.; Suen, C.S.; Chin, W.W.  
J. Biol. Chem. 272, 27629-27634, 1997  
A:Title: TRAM-1, a novel 160-kDa thyroid hormone receptor activator molecule, exhibits c  
A:Reference number: Z15120; MUID:98010595; PMID:9346901  
A:Accession: T03851  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-1424 <TAK>  
A:Cross-references: UNIPROT:Q9Y6Q9; UNIPARC:UPI000012FB45; EMBL:AF016031; NID:G2584879;  
R:Anick, S.L.; Kononen, J.; Walker, R.L.; Azores, D.O.; Tanner, M.M.; Guan, X.Y.; Sautc  
Science 277, 965-968, 1997  
A:Title: AIB1, a steroid receptor coactivator amplified in breast and ovarian Cancer.  
A:Reference number: Z15053; MUID:97400625; PMID:9252329  
A:Accession: T03749  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-1213, 1218-1424 <ANZ>  
A:Cross-references: UNIPARC:UPI000002AEE3; EMBL:AF012108; NID:G2331249; PIDN:AAC51677.1;  
R:Li, H.; Gomes, P.J.; Chen, J.D.  
Proc. Natl. Acad. Sci. U.S.A. 94, 8479-8484, 1997  
A:Title: RAC3, a steroid/nuclear receptor-associated coactivator that is related to SRC  
A:Reference number: Z14950; MUID:97385128; PMID:9238002  
A:Accession: T03443  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-130, 1218-1273, 1277-1424 <LIH>  
A:Cross-references: UNIPARC:UPI0000169657; EMBL:AF010227; NID:G2318005; PIDN:AAC51663.1;  
A:Experimental source: cell line HeLa  
C:Genetics:  
A:Gene: TRAM-1; AIB1; RAC3  
A:Map position: 20

Query Match 6.6%; Score 100.5; DB 2; Length 1424;  
Best Local Similarity 20.2%; Pred. No. 3.8;  
Matches 52; Conservative 44; Mismatches 115; Indels 47; Gaps 11;

Qy 11 TPG-----EFESKYFEHGVRLPPFCRGKMBEIANFPVRPSDVIVTPKSGT 58  
Db 24 TFGQGLTCSGRKRRREQESKYIE---ELABELISANLSDIDNFNVPDK-----C 69  
Qy 59 SLLOEVVYLVSGQADPDEIGLWNIDEQLPVLEYQPQGLDII-KELTSRLIKSHLPYRFL 117  
Db 70 AILKTVRIQRIKEQKG--TISNDDVQKADVSTGGQVIDKOSLGLPLLQALDGFILV 127  
Qy 118 PSDLHNGDSKVIYMARPKDLVSVYQF-----HRSRLTMSYRGTFQFCRRFMNDKLG 171  
Db 128 V---NRDGNIVFVSEN---VTQYLYQKQEDLVNTSVYNILHEEDKDKFLKNLPKSTVN 179  
Qy 172 YGSWFHVQEFWEHRMDSNVLFKLYEDMHRDLVTMVEQLARFLGVSC-DKAOLEALTEHC 230  
Db 180 GVSWTNETQKSHTFNCRMLMKTPHDILEDINASPENRQRYETMQCFALSQPRAMBE 239  
Qy 231 HOLVDQC--CNAEALPVG 246  
Db 240 EDL-QSCMICVARRITTG 256

RESULT 64  
T27665  
hypothetical protein ZK1037.10 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
A:Accession: T27665  
R:Baaham, V.  
submitted to the EMBL Data Library, October 1996  
A:Reference number: Z20401  
A:Accession: T27665  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-481 <WIL>  
A:Cross-references: UNIPROT:O45992; UNIPARC:UPI0000061347; EMBL:Z81142; PIDN:CAB03509.1;  
A:Experimental source: clone ZK1037

C:Genetics:  
A:Gene: CESP:2K1037.10  
A:Map position: 5  
A:Introns: 37/3; 75/3; 141/3; 173/2; 253/1; 298/3; 310/1; 371/2; 424/3

Query Match      6.3%; Score 96; DB 2; Length 481;  
Best Local Similarity 26.9%; Pred. No. 2.1;  
Matches 47; Conservative 25; Mismatches 59; Indels 44; Gaps 11;

Qy    105 PRLIKSHLPYRFLPSDLHNGDSKVIVWARPDKLVVSYQFHRSLRTMSYRGTFQFCRR 164  
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :  
Db    291 PAVAQQPAPYSYLPA---APCQQYYCF--PNDAVVNVIE--KAVKRM----- 331  
  
Qy    165 FMNDKLGYSWFHFVQSFWEHRHSDSNVLFLKYEDMRDLVTMVEQLARFLGVSCDKAOLE 224  
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :  
Db    332 ---DELEIGDWVEALDENGE---DITPLPKYV-LHRD-----PEQAELFLPSLDNGETFF 380  
  
Qy    225 ALTE-----HCQLVDQC-CNAEALPVGRGVLGMKDIFTYSMNSE---KFDLV 268  
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : :  
Db    381 TLTEKLHVYTTCRNSELKISWESISAGKNVAG---DCFYLAQSEALT KYRLV 432

RESULT 65  
B64592  
hypothetical protein HP0578 - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
C:Accession: B64592  
R/Tomb, J.F.; White, O.; Ketzlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, L.; Venter, A.; Zannis-Haber, E.; Adams, J.; Kelley, J.M.; Corton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Son, J.D.; Kelley, J.M., 1997  
Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: B64592  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-678 <TOM>  
A:Cross-references: UNIPROT:Q23301; UNIPARC:UPI00000C0860; GB:AEO00571; GB:AEO00511; NID:

Query Match	6.2%;	Score 94.5;	DB 2;	Length 678;
Best Local Similarity	21.6%;	Pred. No. 4.6;		
Matches	68;	Conservative 49;	Mismatches 113;	Indels 85; Gaps 18
Qy	11	TPGFESKYF-----EFHGVLPPFCGKMEEIA-----NPPVRPS-----DWIIVTPKS 56		
Db	229	TPGAFRNLYLLARNVQRSHNLKFSFAKETPLEVAKNFYHLKENPSNNLYELLTOTSNN 288		
Qy	57	GTSLLQEVYVLVSQA-----DP--DEIGLMNIDQLPVLVEYPOQGLDIIKELTSPLRIK 109		
Db	289	SNQITQHYFYIVSELSLWHFDPKFDATGLTSALQDLVKEHAHMLSAFIE--SAPRTVK 346		
Qy	110	S-----HLPYRFLPSDLHGDSKVIYARNPKDLVSVY-YQFHRSLTMSVGRFTQBF 162		
Db	347	SLDVQITGLPY-----INDNNLV-----NSGVILPSPMAIGNITKTLGKKNFYVGG 394		
Qy	163	RRFMN-----DKLGYGSWF--EHVQEFWEHR-----MDSNVLF---LKVEDMH 200		
Db	395	SGIWNKLTFTKKQGFHALYFNNHLLULEFAQNKPYPKPIESNGWGHNDLFDVILENTNPH 454		
Qy	201	RDLVTMVBQLA-----RFLVGSCKAQALEALTEHCHQLVDGCCNAEALPVGR--GR 249		
Db	455	EKTSVMVMTLSNHAIKNVNLKAFGVPLEKIQ-----QPVEKTPKSENLPDANSLSGH 505		
Qy	250	VGLMKDIIETVSMNEK 264		
Db	506	I-YWYDKVIVSFIKK 519		

RESULT 66  
S18765  
Sd protein - fruit fly (*Drosophila melanogaster*)

C;Species: Drosophila melanogaster  
C;Date: 19-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C;Accession: S18765  
R;Powers, P.A.; Ganetzky, B.  
Genetics 129, 133-144, 1991  
A;Title: On the components of segregation distortion in Drosophila melanogaster. V. Mol  
A;Reference number: S18765; MUID:92038937; PMID:1936954  
A;Accession: S18765  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-363 <POM>  
A;Cross-references: UNIPROT:P25722, UNIPARC:UPI000016BD5D; EMBL:X60218; NID:g11012; PDB:  
C;Genetics:  
A;Gene: FlyBase:Scd  
A;Cross-references: FlyBase:FBcn0024230

[illegible]

RESULT 67

T03387  
polyamine oxidase (EC 1.5.3.11) precursor - maize plasmid pCR2.1  
C:Species: Zea mays (maize)  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C:Accession: T03387  
R:Tavladoraki, P.; Schinina, M.; Cecconi, F.; Di Agostino, S.; Manera, F.; Rea, G.; Mar  
EBBS Lett. 426, 62-66, 1998  
A:Title: Maize polyamine oxidase: primary structure from protein and cDNA sequencing.  
A:Reference number: Z14918; MUID:98258926; PMID:9598979  
A:Accession: T03387  
A:Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-500 <TAV>  
A:Cross-references: UNIPROT:O64411; UNIPARC:UPI00000008E5; EMBL:AJ002204; NID:g3043528;  
A:Experimental source: Cultivar Paolo  
A:Note: Flavin-containing protein  
C:Genetics:  
A:Gene: PAO  
A:Genome: plasmid pCR2.1  
C:Keywords: oxidoreductase

Query Match	6.0%	Score 92.5;	DB 2;	Length 500;
Best Local Similarity	22.5%;	Pred. No. 4.5;		
Matches	56;	Conservative	39;	Mismatches 99; Indels 55; Gaps 11;
Qy	3	ESEAETSTPGCESKSYFER-HGVRLPFFCRGKMEEI--ANFPVPRSDVMWIVTPVKSGTS	59	
Db	176	EHQFNGPATPDMVVYKYDYEAEPRTVSLQNTPLATFSDFGDDVFVAQRGEA	235	
Qy	60	LLEQVYVLVGADPDIGLNMNIDQLPV-----LEYPOGLDIIKE---LTSPRLIKS	110	

Db 236 V---VYLAGQVLKTDGKGIIVPRQLQNKVRIKYPGQVTVKTTEDNSVYSADYVMV 292

Qy 111 HLPVRLPSDLHGDSKVIYARNPKDLVSVYQFHRSLRTMSYRGTFQEFRCRFMNDKL 170

Db 293 SASLGLVQSDL-----IQPKPLPTWKVRAIYQDMAV-----YTKIFLKPRKFWPEGK 342

Qy 171 G-----YGSFHEVQEFWEHRMDSNVLFKYEDMHRDLVTWQOLARFLGVS 217

Db 343 GREPFYASSRGGYYGVW-----QEFKQYPDANVL-----LVTVTDEESRRIRBQQ 388

Qy 218 CD---KAOL 223

Db 389 SDEQTKAEI 397

RESULT 68

T39468

anthranilate synthase (EC 4.1.3.27) component II - fission yeast (Schizosaccharomyces po

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C:Accession: T39468

R:Mc Dougall, R.; Rajandream, M.A.; Barrell, B.G.; Seeger, K.; Quail, M.; Harris, D.

submitted to the EMBL Data Library, July 1999

A:Reference number: Z21856

A:Accession: T39468

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-759 <MCD>

A:Cross-references: UNIPROT:Q92370; UNIPARC:UPI0000013762D; EMBL:AL096874; PIDN:CAB51341.

A:Experimental source: strain 972h-; cosmid c1539

C:Genetics:

A:Gene: SPDB:SPBC1539.09c

A:Map position: 2

C:Superfamily: trpG-trpC-trpF trifunctional enzyme; trpC homology; trpF homology; trpG h

C:Keywords: carbon-carbon lyase; oxo-acid-lyase

Query Match 6.0%; Score 92; DB 2; Length 759;

Best Local Similarity 22.2%; Pred. No. 8.8;

Matches 42; Conservative 30; Mismatches 71; Indels 46; Gaps 7;

Qy 1 MAESEAEPTPT-GEFESKYFEFGVRL-----PPFCRGKMEETANFPVPSDV-W 49

Db 239 MRENSVSVSTKIRKQESILSKIAQRLIDIAESKRKPGLSVGDLTQTVLNINAPPICNF 298

Qy 50 IVTPYKSGTSLQEVVYLVSQADPDEIGLMNIDEQ--LPVLEYPPQGLDIKELTSRL 107

Db 299 YERLKQSKPALMAEV-----KRASPSK-GDIKDANAALQALTYAQVGASVISVITEPKW 352

Qy 108 IKSHL-----PYRFLPSDLHGDSKVIYARNPKDLVV 140

Db 353 FKGSLNDFVARKAVEHVANRPAILRKDFIIDPYQIMEARLNGADSVLLIVAMLSREQLE 412

Qy 141 SYVQFHRSL 149

Db 413 SLYKFSKSL 421

RESULT 69

G71232

hypothetical protein PH0121 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 21-Jul-2000

C:Accession: G71232

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Hino, Y.; Yamamoto, S.; Sekin

M.; Ohtuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a

A:Reference number: A71000; MUID:98344137; PMID:9679194

A:Accession: G71232

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1434 <KAW>

A:Cross-references: UNIPARC:UPI0000015A654; GB:AF000001; NID:g3236128; PIDN:BAA29190.1; H

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by GenBank

C:Genetics:

A:Gene: PH0121

Query Match 6.0%; Score 91.5; DB 2; Length 1434;

Best Local Similarity 22.6%; Pred. No. 23;

Matches 58; Conservative 46; Mismatches 86; Indels 67; Gaps 14;

Qy 45 PSD-VWIVTPYKSGTSLQEVVYLVSQADPDEIGLM-NIDEQLPVLEYPOP----GLDI 98

Db 1095 PEDFVFSLNAAKYHTVIINE--NIVTHQCDGDEDAVMLLLDALLNFSRYILPEKEGG--- 1149

Qy 99 IKELTSRLIKSHLPYRFLPSDLHGND-----SKVIYARNPKDLVVSVYQFHRSL 149

Db 1150 --KMDAPLVITTRLDPREVDSEVHMDIVRYYPLEFYEATYELKSPKELV----- 1197

Qy 150 RTMSYRGTFQFCRFPWNDKLGYSWEHVOEFWEHRMDSNVLFKLYE-DMHRDLVTMVE 208

Db 1198 -----GVIERVEDRLKPEMYIYGLKFTH-----DTDDIALGPQMSLYKQLGDMEE 1242

Qy 209 QLARFLGVSCDKAQLAEALTEHCHQLVDQCCNAEALPVGRGRVGLW--KDIFTVSNNEKF- 265

Db 1243 KVKQLDVA---RRIRAVDE--HKVAETILNSHLIPDLRGNLRSFTTRQEFRCVCNCKR 1297

Qy 266 ----DLVYKQKMGKCDL 278

Db 1298 RPPLD-----GRCPI 1307

RESULT 70

S49879

hypothetical narbonin-like 2S protein - fava bean

C:Species: Vicia faba (fava bean)

C:Date: 05-Mar-1995 #sequence\_revision 14-Jul-1995 #text\_change 09-Jul-2004

C:Accession: S49879

R:Nong, V.; Schlesier, B.; Muentz, K.

submitted to the EMBL Data Library, November 1994

A:Description: The narbonin gene from Vicia faba L.

A:Reference number: S49849

A:Accession: S49879

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-285 <NON>

A:Cross-references: UNIPROT:Q41660; UNIPARC:UPI00000A571C; EMBL:Z46826; NID:g600101; PID

C:Superfamily: alcohol sulfotransferase

Query Match 5.9%; Score 91; DB 2; Length 285;

Best Local Similarity 25.1%; Pred. No. 2.8;

Matches 46; Conservative 25; Mismatches 80; Indels 32; Gaps 9;

Qy 13 GEPESKYF-BPHGVRLPPFCRGKMEET-----ANFPVRPSD--VWIVTPYKSGTS 59

Db 52 GNFEESMWDEFFGDPKVNKLTKHPEVKVVISIGRGVETFPDPAEQNIWVSNVAVKSLK 111

Qy 60 LLQEVVYLVSQADPDEIGL---MNIDEQLPVLEYPPQGLDIKELTSRLIKSHLPYR 115

Db 112 IIQK--YKNESGNLIDGIDINYEHIKSGDEAFPRLLIG-----QLITELKKERDLNIHV-VS 163

Qy 116 FLPSDLHGDSKVIYARNPKDLVVSVYQFHRSLRTMSYRGTFQEFRCRFMNDKLGYSW 175

Db 164 IAPSE--KNASSYLNLYNANPDINLVYQFSNQRLRHVSTEDAFVDIYKRVVND-----Y 216

Qy 176 FEH 178

Db 217 FTH 219

RESULT 71

S49898

hypothetical narbonin-like 2S protein (clone pVSNAG2) - spring vetch

C:Species: Vicia sativa (spring vetch, tare)

C:Date: 05-Mar-1995 #sequence\_revision 14-Jul-1995 #text\_change 09-Jul-2004

C;Accession: S49898  
R;Nong, V.; Muentz, K.  
Submitted to the EMBL Data Library, November 1994  
A;Description: A genomic sequence encoding putative narbonin from *Vicia sativa*.  
A;Reference number: S49880  
A;Accession: S49898  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-285 <N>  
A;Cross-references: UNIPROT:Q41704; UNIPARC:UPI00000A97B2; EMBL:Z46835; NID:G600112; PID  
C;Superfamily: alcohol sulfotransferase

Query Match 5.9%; Score 91; DB 2; Length 285;  
Best Local Similarity 25.1%; Pred. No. 2.8;  
Matches 46; Conservative 25; Mismatches 80; Indels 32; Gaps 9;

Qy 13 GEFESKYP-EPHGVRLPPFCRGKMEI-----ANFVVRPSD--VWIVTPKSGTS 59  
Db 52 GNFESENWDEFFGDPKVKNLTKHPEVKVVISIGRGVETFPDPAEQNIWISNAVKSLKL 111

Qy 60 LLQEVVYLVSGADPDEIGL-----WNIDEQPVLEYPOPGDLIIKELTSPRLIKSHLPYR 115  
Db 112 IIQK--YKNSGNLIDGIDINVEHIKSDAEPRLIG-----QLITELKKERDLNIHV-VS 163

Qy 116 FLPSDLHNGDSKVYMAENPKDLVSVYQFHRSIRLMTSYRGTFQFCFRFNMNDKLYGWSW 175  
Db 164 IAPSE-NNASSYLNDYNANPDINLDVYQFNSQLRHVSTEDAFVDIYKRVND-----Y 216

Qy 176 FEH 178  
Db 217 FTH 219

RESULT 72  
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a  
A;Reference number: A71000; MUID:98344137; PMID:9679194  
A;Accession: A71453  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-508 <KAW>  
A;Cross-references: UNIPROT:O58018; UNIPARC:UPI0000062D81; GB:AP0000001; NID:G3236128; PI  
A;Experimental source: strain OT3  
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C;Genetics:  
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0107

Query Match 5.9%; Score 91; DB 2; Length 508;  
Best Local Similarity 25.5%; Pred. No. 6.2;  
Matches 49; Conservative 26; Mismatches 65; Indels 52; Gaps 10;

Qy 40 NPPVRPSDVWIVTPK-SGTSLLQEVVYLVSGADPDEIGLNM-----IDEQLPV 88  
Db 150 DFPAR-ITAEIVDAPKLSLEGVKKANYLVSEGADIVDIGVAGETNVKFLDVIPEFREV 208

Qy 89 L--EYP-----QFGDLIIKELTS-----PRLIKSHLPYRFLPSDLHNGDSKVYI 130  
Db 209 LGREVPISLDSLNAAKEAGLDFVDMILSDWSNVDELVTDPKPVLIPTDMKNG-----Y 263

Qy 131 MARPKDLVSVYQFHRSIRLMTSYRGTFQFCFRFNMNDKLYGWSWFEHVQFWEHRMDSN 190  
Db 264 PFLNPRDRVSELEKLEKALSIGY-----KRIADLI-----LEHYPGF-----SRSI 306

Qy 191 VLFLKYEDMHRD 202

Db 307 VAFSLYRDRNRE 318

RESULT 73  
T46566  
anthranilate synthase (EC 4.1.3.27), trifunctional - fission yeast (*Schizosaccharomyces*  
N;Contains: glutamine amidotransferase (EC 2.6.1.-); indole-3-glycerol-phosphate synthase  
C;Species: *Schizosaccharomyces pombe*  
C;Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 09-Jul-2004  
C;Accession: T46566  
R;David, C.  
submitted to the EMBL Data Library, October 1996  
A;Reference number: Z23074  
A;Accession: T46566  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-759 <DAV>  
A;Cross-references: UNIPROT:Q92370; UNIPARC:UPI00001691BD; EMBL:Y09137; PIDN:CAA70348.1  
A;Experimental source: strain 972(-)  
C;Genetics:  
A;Gene: trp-1  
A;Map position: II  
A;Superfamily: trpG-trpC-trpF trifunctional enzyme; trpC homology; trpF homology; trpG h  
C;Keywords: aminotransferase; carbon-carbon lyase; carboxy-lyase; intramolecular oxidore

Query Match 5.9%; Score 91; DB 2; Length 759;  
Best Local Similarity 22.2%; Pred. No. 11;  
Matches 42; Conservative 29; Mismatches 72; Indels 46; Gaps 7;

Qy 1 MAESEAEPTSTP-GEFESKYPEFHGVRL-----PPFCRGKMEIANTPPVRPSDV-W 49  
Db 239 MRENSNSVSTKIRKQESILEKHAQRLIDTAESKRKPEGLSVGLQTVLNLNIAPPCINF 298

Qy 50 IVTPYKSGTSLQEVVYLVSGADPDEIGLNMIDEQ--LPVLEYPOPGDLIIKELTSPRL 107  
Db 299 YERLKQSKPALMAEV-----KRASPSK-GDIKLDANAIAQALTVAQVGASVISVLTEPKW 352

Qy 108 IKSHL-----PRFLPSDLHNGDSKVYMAENPKDLVV 140  
Db 353 FKGSNLDFVARKAVEHVANRPAILRKDFIIDPYEIMEARLNGADSVLLIVAMLSREQLE 412

Qy 141 SYQFHRSL 149  
Db 413 SLYFESKSL 421

RESULT 74  
A57169  
[heparan sulfate]-glucosamine N-sulfotransferase (EC 2.8.2.8) - human  
N;Alternate names: N-heparan sulfate sulfotransferase (N-HSST)  
N;Contains: desulfoglycosaminidase (EC 2.8.2.8); glycosaminoglycan N-acetylgluc  
C;Species: *Homo sapiens* (man)  
C;Date: 23-Feb-1996 #sequence\_revision 23-Feb-1996 #text\_change 09-Jul-2004  
C;Accession: A57169; G02129; G01581  
R;Dixon, J.; Loftus, S.K.; Gladwin, A.J.; Scambler, P.J.; Wasmuth, J.J.; Dixon, M.J.  
Genomics 26; 239-244 1995  
A;Title: Cloning of the human heparan sulfate-N-deacetylase/N-sulfotransferase gene from  
A;Reference number: A57169; MUID:95324914; PMID:7601448  
A;Accession: A57169  
A;Molecule type: mRNA  
A;Residues: 1-882 <DIX>  
A;Cross-references: UNIPROT:P52848; UNIPARC:UPI000012CDA5; GB:U18918; NID:G976371; PIDN:  
A;Note: authors translated the codon AAG for residue 42 as Leu, AAA for residue 110 as I  
as Ile  
R;Humphries, D.E.  
submitted to the EMBL Data Library, September 1995  
A;Reference number: G09202  
A;Accession: G02129  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-882 <HOM>  
A;Cross-references: UNIPARC:UPI000012CDA5; EMBL:U36600; NID:G1036796; PIDN:AAC27354.1; P  
R;LaBell, T.L.

submitted to the EMBL Data Library, December 1994  
A:Reference number: G07829  
A:Accession: G01581  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-59, A', 61-688, 'G', 690-742, 'R', 744-882 <LAB>  
A:Cross-references: UNIPARC:UPI000016A159; EMBL:U17970; NID:G841163; PIDN:AAA67765.1; PI  
C:Genetics:  
A:Gene: GDB:HSST; NST1  
A:Cross-references: GDB:593916; OMIM:600853  
A:Map position: 5q32-5q33.1  
C:Function: <DAC>  
A:Description: as glycosaminoglycan N-acetylglucosaminyl N-deacetylase, hydrolyzes the N-  
A:Pathway: heparan sulfate biosynthesis  
C:Function: <NST>  
A:Description: as desulfotransferase, catalyzes the formation of glucosamin  
enosine-5'-phosphosulfate and converting it to adenosine 3',5'-bisphosphate  
A:Pathway: heparan sulfate biosynthesis  
C:Superfamily: Caenorhabditis elegans hypothetical protein F08B4.6  
C:Keywords: glycoprotein; Golgi apparatus; hydrolase; sulfotransferase; transmembrane pr  
F:1-17/Domain: intracellular #status predicted <INT>  
F:18-39/Domain: transmembrane #status predicted <TRM>  
F:40-882/Domain: trans-Golgi network luminal #status predicted <LUM>  
F:231,351,401,667/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.9%; Score 90.5; DB 2; Length 882;  
Best Local Similarity 20.9%; Pred. No. 15;  
Matches 54; Conservative 36; Mismatches 95; Indels 73; Gaps 11;  
QY 12 PGEPSKYFHFHGVLPFPFCRGKMEIANFPVPSDVW-----IVTPKSGT 58  
DB 563 PVQLAQKYFQIFSEKDPQLWQPCD-----KRHKDWSKEKTCDRFPKLLIIGPQKGT 617  
QY 59 SLLQEWVYLVSGADPD-----EIGLMN-----ID---EQLPVLE----- 90  
DB 618 TAL-----VLF-LGMHPDLSNYPSSSETFEEIOFFNGHNYHKIDWMEFFPIPSNTTSD 672  
QY 91 YPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYARNPKDLVVSYYQFHRSL- 149  
DB 673 YFEKSANYFDSEVAPRRAAALLP-----KAKVLTILINPADRAYSWYQORAH 721  
QY 150 RTMSYRGTFQEB-----CRPFMNDKLGYSWFHVEHWRMDSNVLFKLYEDM 199  
DB 722 DPVALKYTFHEVITAGSDASSKLALQNRCLVPGWYATHIERKWSAYHANQILVLDGKLL 781  
QY 200 HRDLVTMVEQLARFLGVS 217  
DB 782 RTEPAKVMWQKFLGVT 799

RESULT 75  
D72379  
hypothetical protein TM0428 - Thermotoga maritima (strain MSB8)  
C:Species: Thermotoga maritima  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C:Accession: D72379  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Swinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;  
C.M.  
Nucle 399, 323-329, 1999  
A>Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq  
A:Reference number: A72200; MUID:99287316; PMID:10360571  
A:Accession: D72379  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-435 <ARN>  
A:Cross-references: UNIPROT:Q9WYQ9; UNIPARC:UPI00000D3A0F; GB:AE001721; GB:AE000512; NID  
A:Experimental source: strain MSB8  
C:Genetics:  
A:Gene: TM0428  
C:Superfamily: NADPH dehydrogenase chain OYE2

Query Match 5.9%; Score 90; DB 2; Length 435;

Best Local Similarity 21.0%; Pred. No. 6.1;  
Matches 44; Conservative 42; Mismatches 64; Indels 60; Gaps 11;  
QY 58 TSLQEVVYLVSGADPDPEIGLMNIDEOLP-----VLEYPOP-----GL 96  
DB 226 TKLLKDIIVLRKVEIDITVRLNLSDFIPIYPYWGSGTEEGTLDTFEPCRLKKEELGV 285  
QY 97 DIIKELTSPRLIKSHL--PY-----RFLPSDLHN--GDSKVIYARNPKD-----LVVSY 142  
DB 286 KIINVSTASTPYLKPHINRPYDEMGKYNPPE-HPIVGMARLLNLAKLAKETSKTLVWASG 344  
QY 143 YOFHSLRTMSYRGTFQEB-FCRRFMNDKLGYSWFHVEHWRMDSNVLFKLYEDMHR 201  
DB 345 FTWFRQFAPYVAAGMLKNGWC-----DFVGFGR-----MTFAIPDPK 382  
QY 202 DLVTMVEQLARFLGVSCKD-KAQLKALTEHC 230  
DB 393 DILIKGELDPKKVCITCNKCAELKAAGESC 412

RESULT 76  
B89791  
hypothetical protein SA0261 [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A>Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A:Reference number: A89758; MUID:21311952; PMID:11418146  
A:Accession: B89791  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-332 <KUP>  
A:Cross-references: UNIPROT:Q99WV4; UNIPARC:UPI0000054670; GB:BA000018; PID:gl3700187;  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: SA0261  
C:Superfamily: lac repressor

Query Match 5.8%; Score 89; DB 2; Length 332;  
Best Local Similarity 25.4%; Pred. No. 5.2;  
Matches 36; Conservative 16; Mismatches 38; Indels 52; Gaps 6;  
QY 24 GVRLP----PFCRGKMEIANFPVPSDVMTVTPKSGTSLLEQVYLVSGADPDPEIGL 79  
DB 66 GVILPSLTNPFPSALMQSIHDH--KPSDVIDLCFLTSTATDLYDNIKHLIDRID----- 117  
QY 80 MNIDEQLPVLEYPOPGLDIIKELTSP-----RLIKSHLPYRFLPSDLHNGDSKVIYARN 134  
DB 118 -----GLIIAQYISSPDALNNYKHHVYVYVLDQNDHOG----- 152  
QY 135 PKDLVVSYYQFHRSLRTMSYRG 156  
DB 153 -----YTDF---VRTNEYQG 164

RESULT 77  
A42855  
N-heparan sulfate sulfotransferase - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Jun-2000  
C:Accession: A42855  
R:Hashimoto, Y.; Orellana, A.; Gil, G.; Hirschberg, C.B.  
J. Biol. Chem. 267, 15744-15750, 1992  
A>Title: Molecular cloning and expression of rat liver N-heparan sulfate sulfotransferas  
A:Reference number: A42855; MUID:92348437; PMID:1379236  
A:Accession: A42855  
A:Status: preliminary  
A:Molecule type: mRNA; protein  
A:Residues: 1-882 <HAS>





Db 159 EIHQVLL-----QVADPPEKVASVIDHTYPN-----KQVVQLISSQIDADRM DYL 204

Qy 81 NIDEQLPVLEYPQGLDIIKELTSRLI:KSHLPYRFLPSDLHNGDSKVIYMARPKDLVV 140

Db 205 LRDSYFTGASYGE--FDLTRILRVIRPIENGIAFO-----RNGMHAI-----EDYVL 249

Qy 141 SYVO-----FHRSLRTWSYRGTFQFCRR-----FMNDKLGYSWFHVFQEFWEHRMDSN 190

Db 250 SRYQMYMQVYFHPATRAHEV--LLQNLLKRAKELYPEDKDFARTSPHLLFFFEKN----- 303

Qy 191 VLFLKYEDMHRDLVTMVQLARFLGVSCDKAQL-----EALTECHQOLV----- 234

Db 304 -----VLTLDYALDDGVMTYFQLWMTSPDKILADLSHFRVNRKVPKSTTF 350

Qy 235 -----DQCCNAEALPVGRGVGLWKDIFTVSMNEKFDLIVY 269

Db 351 SQEODQQLASMRKLVED---IGFDDPYDT-AIHKNFIDLFIY 386

RESULT 84

A49733

[heparan sulfate]-glucosamine N-sulfotransferase (EC 2.8.2.8) - mouse

N:Alternate names: N-heparan sulfate sulfotransferase (N-HSST)

N:Contains: desulfoheparin sulfotransferase (EC 2.8.2.8); glycosaminoglycan N-acetylgluc

C:Species: Mus musculus (house mouse)

C>Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 09-Jul-2004

C:Accession: A49733; A53736

R:Orellana, A.; Hirschberg, C.B.; Wei, Z.; Swiedler, S.J.; Ishihara, M.

J. Biol. Chem. 269, 2270-2276, 1994

A:Title: Molecular cloning and expression of a glycosaminoglycan N-acetylglucosaminyl N-

A:Reference number: A49733; MUID:94124588; PMID:8294485

A:Accession: A49733

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-883 <ER>

A:Cross-references: UNIPROT:PS2850; UNIPARC:UPI0000000997; GB:U02304; NID:9457943; PIDN:

R:Eriksson, I.; Sandbaeck, D.; Ek, B.; Lindahl, U.; Kjellen, L.

J. Biol. Chem. 269, 10438-10443, 1994

A:Title: cDNA cloning and sequencing of mouse mastocytoma glucosaminyl N-deacetylase/N-

A:Reference number: A53736; MUID:94193735; PMID:8144627

A:Accession: A53736

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-117, 'V', 119-395, 397-883 <ERI>

A:Cross-references: UNIPARC:UPI0000163857; GB:X75885; NID:9474430; PIDN:CAA53479.1; PID:

C:Superfamily: Caenorhabditis elegans hypothetical protein F08B4.6

C:Keywords: carboxylic ester hydrolase; glycoprotein; sulfotransferase; transmembrane pr

Query Match 5.7%; Score 87.5; DB 2; Length 883;

Best Local Similarity 20.2%; Pred. No. 27; Mismatches 109; Indels 53; Gaps 9;

Matches 50; Conservative 36;

Qy 12 PQEFESKVFEPHGVRLPFPFCRGKMBEIANFPVRPSDVW-----IVTYPKSGT 58

Db 562 VPVLAQKYFELPQPERSPLWNPCCD-----KRHKDWSKENTCDRLPKFLVIGPKGTG 616

Qy 59 SLLQEVVYL-----VSCGADP---DBIGLMN-----IDEQLPVLEYP-QPGLDIKELT 103

Db 617 TAIHFPLSLHPAVTSFPSPSTFEBIQFNGPNYHKGDWYMDFFPVPNSASTDFLFEKS 676

Qy 104 SPRLIKSHLPYR--FLPSDLHNGDSKVIYMARPKDLVSYVQPHRS-----L 149

Db 677 ATYFDSEVVPVRGAALLPR-----AKITVLINPADRAYSWYQHRAGDPDIALNYTFY 730

Qy 150 RTMSYRGTFQFCRRFMNDKLGYSWFHVFQEFWEHRMDSNVLFLKYEDMHRDLVTMVEQ 209

Db 731 QVISASSQAPLLRLSLQNRCLVPGVGYSTHLQRWLTYTPSGQLLIMDGQELRWPAASMEI 790

Qy 210 LARFLGVS 217

Db 791 IQKFLGIT 798

RESULT 85

C84500

probable retroelement pol polyprotein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: C84500

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: C84500

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-976 <STO>

A:Cross-references: UNIPARC:UPI000017A748; GB:AB002093; NID:94557068; PIDN:AAD22507.1; G

C:Genetics:

A:Gene: At2g12020

A:Map position: 2

Query Match 5.7%; Score 87.5; DB 2; Length 976;

Best Local Similarity 25.0%; Pred. No. 30;

Matches 47; Conservative 36; Mismatches 80; Indels 25; Gaps 12;

Qy 83 DEQLPVLEYPQGLDII-KELTSRLI--KSHLPYRFL-----PSDLHNGDSKVIYMARN 134

Db 8 DESKSSVEH-QRRNLNIQKRLRLKLLSYWKQVLSIRPQTAVGPDVHRLOKABCCYQER 66

Qy 135 PKDLVWSYVQPHRSLRTWS--YRGTFQFCRRFMNDKLGYSWFHVFQEFWE--HRMDS 189

Db 67 PFPLTI--YRSHVGERGMSIFTDMIEDIMEVFMNDFSIVGSLFEDCLNLYKVLARCEE 124

Qy 190 NVLFKYEDMH---RDLVTWVEQLARFLGVSCDKAQLTEHCHQLVDQCCNAEALPVG 246

Db 125 KHLVLNNEKCHFRVQDGVILGHRISEY-GIEADRAKIEVMTS--LQALD---NLKAVRSF 178

Qy 247 RGRVGLWK 254

Db 179 LGHAGFYR 186

RESULT 86

E75199

DNA-directed DNA polymerase (EC 2.7.7.7) II large chain PAB2404 - Pyrococcus abyssi (str

C:Species: Pyrococcus abyssi

C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 05-Oct-2004

C:Accession: E75199

R:anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

A:Reference number: A75001

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1455 <KAW>

A:Cross-references: UNIPROT:Q9V2F4; UNIPARC:UPI0000003472A; GB:AJ248283; GB:AL096836; NID

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB2404

C:Superfamily: DNA polymerase II large subunit

C:Keywords: nucleotidyltransferase

Query Match 5.7%; Score 87.5; DB 2; Length 1455;

Best Local Similarity 22.9%; Pred. No. 52;

Matches 57; Conservative 46; Mismatches 89; Indels 57; Gaps 13;

Qy 37 EIANFP-VRPSD--VWIVTYPKSGTSLQLQEVVVLVSQADPDEIGLM-NIDEQLPVLEYP 92

Db 1104 EIKNIKVVRANDDFVSLNAKKYHNVIINE--NIWTHCCDGEDAVMLLLDALLNFSRY 1161

Qy 93 QP-----GLDIIKELTSRLIKSHLPYRFLPSDLHNGD-----SKVIYMARPKDLV 139



Db 1162 LPEKRG-----KMDAPLVITTRLDPREVDSEVHNMDIVRYYPLEPYEATYELKSPKELV 1216

QY 140 VSYIQPHRSRLTMSYRGTFQEFRCRFEMNDKLGYSWFHEHVQBFWHRMDSNVLFKYE-D 198

Db 1217 -----GVIERVEDRLGKPEMYGLKFTH-----DTDDIALGPKMS 1251

QY 199 MHRDLVTWVEQLARFLGVSCDRAQLEALTEHCHQLVDQCCNAEALPVGRGRVGLW--KDI 256

Db 1252 LYKQLGDMEEKVRROLEVA---KRIRAVDE--HGVAEKILNGLHLPDLRGNLSRFTROBF 1306

QY 257 FTVSMNEKF 265

Db 1307 RCVKCNTKF 1315

RESULT 87

A95150

conserved hypothetical protein SP1290 [imported] - Streptococcus pneumoniae (strain TIGR

C:Species: Streptococcus pneumoniae

C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004

C:Accession: A95150

R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,

nson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: A95150

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-434 <KUR>

A:Cross-references: UNIPROT:Q97QC9; UNIPARC:UPI00000517B0; GB:AE005672; PIDN:AAK75394.1;

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP1290

C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ1154

Query Match 5.7%; Score 86.5; DB 2; Length 434;

Best Local Similarity 21.8%; Pred. No. 12;

Matches 61; Conservative 38; Mismatches 98; Indels 83; Gaps 13;

QY 21 EPHGVRLPPFCRGKMEETIANFPVRSDVWIVTPKSGTSLLOEVVVLVSQGADPEIGLM 80

Db 127 EIHQVLL-----QVAPDFPEKVASVIDHTYPN-----KQVQLISSQIDADRMDYL 172

QY 81 NIDEQLPVLEYPPQGLDIITKELTSRLIKSHLPYRFLPSDLNGDSKVIYMARNPCKDLVV 140

Db 173 LRDSYFTGASYGE--FDLTRIILRVIRPIENGIAFQ-----RNGMHAI-----EDYVL 217

QY 141 SYIQ-----FHSRLTMSYRGTFQEFRCR-----FMNDKLGYSWFHEHVQBFWHRMDSN 190

Db 218 SRQYMYQVYFHPATRAVEV--LLQNLLKRAKELYPEDKDFARTSPHLLPFPEKN---- 271

QY 191 VLFLKYEDMHRDLVTWVEQLARFLGVSCKAQL-----EALTEHCHQLV----- 234

Db 272 -----VTLDYDALDDGVMTYFQLWMTSPDKILADLSHRFVNRKVKFSITF 318

QY 235 -----DOCCNAEALPVGRGRVGLWKDIFTVSMNEKFDLVY 269

Db 319 SOEDQDQLTSMRKLVED----IGFDPDYIT-AIHKNFDLPY 354

RESULT 88

JDVLD

DNA-directed DNA polymerase (EC 2.7.7.7) - duck hepatitis virus

C:Species: duck hepatitis virus, DHBV

C:Date: 20-Sep-1984 #sequence\_revision 08-Nov-1996 #text\_change 09-Jul-2004

C:Accession: S12844; S36614; A00710

R:Mattes, F.; Tong, S.; Teubner, K.; Blum, H.E.

Nucleic Acids Res. 18, 6140, 1990

A:Title: Complete nucleotide sequence of a German duck hepatitis B virus.

A:Reference number: S12844; MUID:91045092; PMID:2235507

A:Accession: S12844

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-836 <MAT>

A:Cross-references: UNIPROT:Q66400; UNIPARC:UPI00001726E5; EMBL:X12798; NID:G59057

A:Experimental source: isolate DHBV FL-6

A>Note: this ORF is not annotated in GenBank entry DHBVFL6, release 103

R:Munshi, A.; Panda, S.K.

A:Description: Cloning sequencing and sequence comparison of the indian isolate.

A:Reference number: S36614

A:Accession: S36614

A:Molecule type: DNA

A:Residues: 1-35,'V', 37-81,'I', 83-115,'Y', 117-176,'Y', 178-184,'T', 186-188,'Y', 190-220,'I'

A:Cross-references: UNIPARC:UPI00000F76D8; EMBL:X74623; NID:G9397342; PIDN:CAAS2700.1; PI

A:Experimental source: isolate IDHBV

R:Mandart, E.; Kay, A.; Galibert, F.

J. Virol. 49, 782-792, 1984

A:Title: Nucleotide sequence of a cloned duck hepatitis B virus genome: comparison with

A:Reference number: A92997; MUID:84138772; PMID:6699938

A:Accession: A00710

A:Molecule type: DNA

A:Residues: 390-734,'A', 736-836 <MAN>

A:Cross-references: UNIPARC:UPI00001726E6; GB:K01834

A>Note: only part of the sequence reported in the GenBank entry is shown in the publishe

C:Superfamily: hepatitis virus DNA-directed DNA polymerase

C:Keywords: DNA biosynthesis; nucleotidyltransferase

Query Match 5.6%; Score 86; DB 1; Length 836;

Best Local Similarity 19.3%; Pred. No. 33;

Matches 68; Conservative 55; Mismatches 119; Indels 110; Gaps 18;

QY 3 ESEATPTGPEFEFSKYEFHFGVLRPPFCRGKMEETIANFPVRPSDVWIVTPKSGTSLIQ 62

Db 383 EAATRGRTSGKSVSPR---DSSAIPVRTSGASDK--NSPLEEENVVLRGNTSPENRIT 437

QY 63 EVVYLVSOGADPEIGLMNIDQLPVLEYPO-----PGLDIIEKLT---S 104

Db 438 GKFLVLDKNS-----RNTPEARLVDFQSFGKNAMRFPYVWSPNLSLTRLRLPVGM 490

QY 105 PRLIKS-----HLPY-----RFLPSDLNGDSKVIYMARNPCKDLVSYVQFHRSLRTM 152

Db 491 PRISLDLSQAYHULPLNPASSRLAVSD---GQRVYFTRKAPMGVGLSPFLHLHFTTAL 546

QY 153 SYRGTFQFCRRF-----MND-----KLGYSWFHEHVQ--- 180

Db 547 G-----SEISRRFNVTFTYMDDFLLCHPNARHLNAISHAVCSFLQELGIRINPDKTTPS 601

QY 181 -----EFWEHRMDSNVLFKYEDMH--RDLVTWVEQLARFLGVSCKAQLTEHCHQLV 234

Db 602 PVNEIRFLGYQIDEN--FMKIEESRWKELRTVIKKI--VGEWYDWKCIQRFVGHNFV- 656

QY 235 DOCCNAEALPVGRGRVGLWKDIF---TVSMNEKFD---LVYKQKMGKCDL 278

Db 657 -----LPFTKGNIELMKPMYAAITNQVNFSSFSSTYRTLLYKLTMGVCKL 700

RESULT 89

T29107

polymorphic antigen p150 precursor - Theileria parva

C:Species: Theileria parva

C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T29107

R:Skilton, R.; Iams, K.P.; Macklin, M.D.; Gobright, E.

submitted to the EMBL Data Library, September 1995

A:Description: Characterisation of a polymorphic 150 kilodalton antigen of Theileria pa

A:Reference number: Z20571

A:Accession: T29107

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1452 <SKI>

A:Cross-references: UNIPROT:Q27028; UNIPARC:UPI0000074921; EMBL:L47230; NID:G986990; PID

A:Experimental source: strain Muguga; ssp. parva



Db 184 ILPED-WHLRFP-SGSEIIQYAAASHVKNKSLDDEQLLDRRRVEYDIFLLVBLHVLVDII 241

QY 93 QPGLDIIKEL-----TSPRLIKSHLPYRFLPSDL-HNGDSKVIYMARPNKDL 138

Db 242 RKGFGSVDEFIALANSVNRKSRACKSLLEHLHLFIEHGLRHFPATQAITGCKKPDFL 301

QY 139 VVSYYOFH-----RSRLTWSVRGTQFECRRPMNDKLGYSWFHVEHWRMDSNVL 192

Db 302 PFSAGAYHDTFPEVNLMLAVKTTCKDRWRQILNEA-----DKIHQVHLFTLQEGVS 354

QY 193 FLKYEDMRDLVTMVEQLARFLGVSCKAQLEALTEHCHQVLDVOCNNAALPVG---RCR 249

Db 355 LAQYREMRSGRLV-----VPSSLHKKYPEAVRAELMLTGAFIAEL 396

QY 250 VGLWKDI 256

Db 397 TGLYADI 403

RESULT 93

S75889

hypothetical protein slr1173 - *Synechocystis* sp. (strain PCC 6803)

A:Accession: S75889

A:Species: *Synechocystis* sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* s.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S75889

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-463 <KAN>

A:Cross-references: UNIPROT:P74254; UNIPARC:UPI000000C0E3E; EMBL:D90913; GB:AB001339; NID F:33-494/Domain: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 5.6%; Score 85; DB 2; Length 463;

Best Local Similarity 21.7%; Pred. No. 18;

Matches 65; Conservative 37; Mismatches 87; Indels 110; Gaps 17;

QY 15 PFSKYFEFHGR-----LPPFCRGKMBEIAN-----FPVRPSDVMIVTPKS 56

Db 116 FDNNPCETYGGRQFMDAPAPSMTRGLQEKISNEALGVTIENQFSVGEYDILLISAKES 175

QY 57 G-TSLQEVVLVSGADPDEI-----GLMNIQDLVLYPQPGLDIIKEL-----TS 104

Db 176 NGLTWNQNNYRIPPGA-TDVLGAYIKQGLKFFVAKVNLKBFDRQGFQALRPLMAYES 234

QY 105 PRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSRLTMSYR----- 155

Db 235 PRFM---LPRL---GMWNAD-----GPELIVLLSPQGAVENTYRTEKIPSNLD 280

QY 156 -----GTQFECRRPMNDKLGYSWFHVEHWRMDSNVLFLKYE-DMHRDLVTMVE 208

Db 281 LPEFVQGEFGQF-----YGAMFDTAYK-----RSGKNVAFLEYAMDVG----- 318

QY 209 QLARFLGVSCKAQLEALTEHCHQVLDVOCNNAAL-PVGRGRVGL-WKDIFTVSNNKFP 265

Db 319 -----SCDP-----CSADPLSPQLEBAGVFWLDPQSPNPSPF 352

RESULT 94

JQ1535

hemagglutinin - phocine distemper virus (strain Ulster/88)

N:Alternate names: attachment protein

C:Species: phocine distemper virus

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004

C:Accession: JQ1535

R:Curran, M.D.; O'Loan, D.; Kennedy, S.; Rima, B.K. J. Gen. Virol. 73, 1189-1194, 1992

A:Title: Molecular characterization of phocine distemper virus: Gene order and sequence

A:Reference number: JQ1535; MUID:92268877; PMID:1588321

A:Accession: JQ1535

A:Molecule type: mRNA

A:Residues: 1-607 <CUR>

A:Cross-references: UNIPROT:P28882; UNIPARC:UPI0000174A8E; GB:D10371

C:Genetics:

A:Gene: H

C:Superfamily: measles virus hemagglutinin

C:Keywords: glycoprotein; hemagglutinin; transmembrane protein

F:39-55/Domain: transmembrane #status predicted <TM>

F:19,149,276,391,422,456,587/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.6%; Score 85; DB 1; Length 607;

Best Local Similarity 21.9%; Pred. No. 26;

Matches 40; Conservative 29; Mismatches 76; Indels 38; Gaps 8;

QY 49 WIVTPYKSGTSLQEVVLVSGADPDEIGL-----MNIDEQLPVLEYQPGLD 97

Db 449 WLTIPTPKNGT-----ILGLINQASKGDFIVTTPHILTFAPRESSTDCHLPIQTYIQDDDD 503

QY 98 IIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDLVVSYYQFHRSRLTMSYRGT 157

Db 504 VLLESNLVVLPTQSEYVVVATYDVSERDHAIVVYVVDARTVSYTYPPR--LRTKGRPDI 561

QY 158 FOEFCRRPMNDKLGYSWFHVEHWRMDSNVLFLKYEDMHRDLVTMVEQLARFLGVS 217

Db 562 LRIEC-FV-----WDGHL---WCHQ-----FYRFQLDATNSTSVENLIR-IRFS 601

QY 218 CDK 220

Db 602 CDR 604

RESULT 95

S76557

carbamoyl-phosphate synthase (glutamine-hydrolyzing) (EC 6.3.5.5) large chain [similarit

C:Species: *Synechocystis* sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Sep-2000

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* s.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S76557

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1105 <KAN>

A:Cross-references: UNIPARC:UPI0000164C78; EMBL:D64002; GB:AB001339; NID:G1001612; PIDN F:584-1024/Domain: biotin carboxylase homology <BC1>

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Start codon: GTG

C:Superfamily: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain; biotin

C:Keywords: ligase

F:33-1086/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain homol

F:33-494/Domain: biotin carboxylase homology <BC1>

F:584-1024/Domain: biotin carboxylase homology <BC2>

Query Match 5.6%; Score 85; DB 2; Length 1105;

Best Local Similarity 23.6%; Pred. No. 59;

Matches 61; Conservative 32; Mismatches 98; Indels 68; Gaps 13;

QY 1 MAESEATPSTPGFEFSK-YFEFHGVRLLPPRCRGKMBEIANFPVRPSDVMIVTPKSGTS 59

Db 621 MVNSNPETVSTDYDTSRLYFE-----PLTK---EDVLNIIAEANPVGIIQFGGQTP 670

QY 60 L-----LQEVVLVVSQ-----GADPDSEIGLMNIDEQ---LPLVLEYQPGLDIIKE 101

Db 671 LKLAIVPLQK--YLNASPCDPVQTKWGTSIDTAEDRERFEKILHELEISQPPNGIARD 728

```
QY 102 LTSPLIKSHLPYRFL--PSDLHNGDSKVIYMARPKDLVSVYQFHRSLRTMSYRGTFQ 159
Db 729 YEESRVANRISYPVVRPSYVLGRAMEI-----VYSDELERVMTAVQIE 776
QY 160 E-----FCRRFMND-----KLGYGSWFHVGQEFWEHRMDSNVLFKYMHR 201
Db 777 PDHPILIDKFLNAIEVDVDSLTDTGKGVIGSIMEHIEAGIHGDS-ACSIPTYSLSD 835
QY 202 DLVTWV-----EQARFLGV 216
Db 836 NVLTTRIQWTEQARALNV 854

RESULT 96
D84514
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 31-Dec-2004
C:Accession: D84514
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84514
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-782 <STO>
A:Cross-references: UNIPROT:Q9S147; UNIPARC:UPI0000178828; GB:AE002093; NID:G4587686; PI
C:Genetics:
A:Gene: At2g14120
A:Map position: 2

Query Match 5.5%; Score 84.5; DB 2; Length 782;
Best Local Similarity 25.2%; Pred. No. 41;
Matches 52; Conservative 20; Mismatches 65; Indels 69; Gaps 10;

QY 3 ESEATPSTPGE-----FESKYFEFHGVRLPPFCRGKMERIANPPV--RPSD 47
Db 121 ETEAETNRVSGNKGVS DIPIGLKIFSPNVLDISLVDLP-----GITKVPVGDQPSD 172
QY 48 V-----WTVTPK-----SGTSLQEVVVLVSQADPD---EIGLWNIDEQLP 87
Db 173 IEARTRTWLTITIKPSCILLAVSPANTDLANSDALQTAGNADPDGHTIGVIT----- 226
QY 88 VLEYPPQGLDIKEILTSPR--LIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVSVYQF 145
Db 227 -----KLDIMDRGTDARNHLLGKTIPLRLGYGVGVNRSQEDILMRSIKDALVAEKF 279
QY 146 HRLSLRTMSYRGTFQFCRRFMNDKLG 171
Db 280 FRSRPV--YSG-----LTDRLG 294

RESULT 97
B49132
fat facets (faf) splice form 1 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 19-Dec-1993 #sequence_revision 25-Apr-1997 #text_change 01-Dec-2000
C:Accession: B49132; A49132
R:Fischer-Vize, J.A.; Rubin, G.M.; Lehmann, R.
Development 116, 985-1000, 1992
A:Title: The fat facets gene is required for Drosophila eye and embryo development.
A:Reference number: A49132; MUID:93202020; PMID:1295747
A:Contents: isogenic st
A:Accession: B49132
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-2747 <FIS>
A:Cross-references: UNIPARC:UPI000016BBEA; GB:L04959; NID:gl57411; PIDN:AAF01345.1; PID:
A>Note: sequence inconsistent with the nucleotide translation
A>Note: sequence extracted from NCBI backbone (NCBIN:127836, NCBIN:129008, NCBIIP:127839)
```

```
A:Accession: A49132
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-2704, 'VT', 2707, 'ANNV', <FI2>
A:Cross-references: UNIPARC:UPI000016BBE9; GB:L04958; NID:gl57410; PIDN:AAF01346.1; PID:
A>Note: sequence extracted from NCBI backbone (NCBIN:127836, NCBIN:129008, NCBIIP:129029)
C:Keywords: alternative splicing

Query Match 5.5%; Score 84; DB 2; Length 2747;
Best Local Similarity 23.7%; Pred. No. 2.5e+02;
Matches 61; Conservative 37; Mismatches 65; Indels 94; Gaps 18;

QY 10 STPGEFESKYFEFHGVRLP-----PFCRGKMERIA-NPPVRPSPDVWIVT-----YPKSGTS 59
Db 1386 NSSGDFEDE-----VIVPDQDFSMCKEALLEVITISFILNPSANEALTSPPNPKFTTS 1439
QY 60 LLQEVVVLVSQADPDDEIGLWNIDEQLPVLEYPPQGLDIIKELTSPRLIKSHLPYRFLPS 119
Db 1440 I-----VLKNP---LRHVQRVASEQL-----FLAS 1461
QY 120 DLHNGDSK-VIYMARPKDLVSVYQFHRSLRTM--SYRGTFQEF-----CRRFMNDKLG 172
Db 1462 TYCAGDRRPFYVW-----NLLVG-----ALKTLLVPQVESTCAEFFSVLCR-----TLSY 1506
QY 173 G---SWPEHVOE-----FWEHRMDSNVLFKYMHRDL-----VTMVEQLARFLGVSC 218
Db 1507 GCYINWPLQISEGLLGDEIKWLQIRENVHATGTQVHEELLEGLHCLAKSLMFLGAD- 1565
QY 219 DKAQLEALTECHQIOLD 235
Db 1566 SKAQLNEL---IHOLID 1579

RESULT 98
C81744
ATP-dependent Clp proteinase, ATP-binding regulatory chain ClpX TC0078 [imported] - Chla
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 05-Oct-2004
C:Accession: C81744
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: C81744
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-419 <TET>
A:Cross-references: UNIPROT:Q9PLM1; UNIPARC:UPI00000577B4; GB:AE002275; GB:AE002160; NID
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
A:Gene: TC0078
C:Superfamily: ATP-dependent protease ATP-binding subunit; FtsH/SEC18/CDC48-type ATP-bin

Query Match 5.5%; Score 83.5; DB 2; Length 419;
Best Local Similarity 23.4%; Pred. No. 21;
Matches 49; Conservative 22; Mismatches 61; Indels 77; Gaps 10;

QY 30 FCRGKMERIANPPVRPSPDVWIVT-----PKSGTSLQEVVVLVSQADPDDEIGLM 80
Db 11 FCGRSEKDVKLIAGPS-VVICDYCIKLCSGILDKTPTPAAGV--STSQA----- 59
QY 81 NIDEQLPVLEYPPQGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPKDLVV 140
Db 60 -----POTSLKVL---TPKEIKRHI-----DSYVVGQERAKKTISV 92
QY 141 SVYQFHRSLRTMSYRGTFQFCRRFMNDK-LGYGSWFHVGQEFWEHRMDSNVLFKYMEDM 199
Db 93 AVYNYHKIRAL-----MQDKQVSYGK-----SNVLLLGPTGS 125
QY 200 HRDLVTMVEQLARFLGVSCDKAQLEALTE 228
Db 126 GKTLI--AKTLAKILDVPPFTIADATTLTE 152
```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.8  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.  
OM protein - protein search, using sw model  
Run on: May 17, 2006, 11:05:06 ; Search time 51 Seconds  
(without alignments)  
487.425 Million cell updates/sec  
Title: US-10-768-158-2  
Perfect score: 1530  
Sequence: 1 MAESEAEPTSPGPESEKYF.....FDLVYKQKMGKDLTFDFYL 284  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 650591 seqs, 87530628 residues  
Total number of hits satisfying chosen parameters: 650591  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries  
Database : Issued Patents AA: \*  
1: /EMC\_Celerra\_IDS33/ptodata/2/iaa/5\_COMB.pep:\*  
2: /EMC\_Celerra\_IDS33/ptodata/2/iaa/6\_COMB.pep:\*  
3: /EMC\_Celerra\_IDS33/ptodata/2/iaa/7\_COMB.pep:\*  
4: /EMC\_Celerra\_IDS33/ptodata/2/iaa/H\_COMB.pep:\*  
5: /EMC\_Celerra\_IDS33/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /EMC\_Celerra\_IDS33/ptodata/2/iaa/RE\_COMB.pep:\*  
7: /EMC\_Celerra\_IDS33/ptodata/2/iaa/backfiles.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1530	100.0	284	US-09-786-240-7 Sequence 7, Appli
2	472	30.8	295	US-09-854-122-21 Sequence 21, Appli
3	456	29.8	295	US-09-328-174A-3 Sequence 3, Appli
4	449	29.3	296	US-09-949-016-6018 Sequence 6018, Ap
5	449	29.3	296	US-09-792-695A-2 Sequence 2, Appli
6	435	28.4	309	US-09-949-016-10328 Sequence 10328, A
7	434	28.4	294	US-08-325-562-2 Sequence 2, Appli
8	434	28.4	294	US-08-437-795-2 Sequence 2, Appli
9	430	28.1	295	US-09-150-133-13 Sequence 13, Appli
10	430	28.1	295	US-09-150-141-13 Sequence 13, Appli
11	430	28.1	295	US-09-374-493-13 Sequence 13, Appli
12	430	28.1	295	US-09-374-824-13 Sequence 13, Appli
13	430	28.1	295	US-09-374-492-13 Sequence 13, Appli
14	430	28.1	295	US-09-785-343-13 Sequence 13, Appli
15	430	28.1	295	US-10-411-976-13 Sequence 13, Appli
16	430	28.1	295	US-10-201-525-13 Sequence 13, Appli
17	422	27.6	304	US-09-609-816-5 Sequence 5, Appli
18	422	27.6	304	US-10-199-334-5 Sequence 5, Appli
19	422	27.6	304	US-10-199-329-5 Sequence 5, Appli
20	422	27.6	304	US-10-199-330-5 Sequence 5, Appli
21	419.5	27.4	283	US-09-609-816-8 Sequence 8, Appli
22	419.5	27.4	283	US-09-609-816-9 Sequence 9, Appli
23	419.5	27.4	283	US-10-199-334-8 Sequence 8, Appli
24	419.5	27.4	283	US-10-199-334-9 Sequence 9, Appli
25	419.5	27.4	283	US-10-199-329-8 Sequence 8, Appli
26	419.5	27.4	283	US-10-199-329-9 Sequence 9, Appli

27	419.5	27.4	283	US-10-199-330-8	Sequence 8, Appli
28	419.5	27.4	283	US-10-199-330-9	Sequence 9, Appli
29	408.5	26.7	283	US-09-609-816-10	Sequence 10, Appli
30	408.5	26.7	283	US-10-199-334-10	Sequence 10, Appli
31	408.5	26.7	283	US-10-199-329-10	Sequence 10, Appli
32	408.5	26.7	283	US-10-199-330-10	Sequence 10, Appli
33	407	26.6	304	US-09-609-816-6	Sequence 6, Appli
34	407	26.6	304	US-10-199-334-6	Sequence 6, Appli
35	407	26.6	304	US-10-199-329-6	Sequence 6, Appli
36	407	26.6	304	US-10-199-330-6	Sequence 6, Appli
37	401	26.2	304	US-09-609-816-4	Sequence 4, Appli
38	401	26.2	304	US-10-199-334-4	Sequence 4, Appli
39	401	26.2	304	US-10-199-329-4	Sequence 4, Appli
40	401	26.2	304	US-10-199-330-4	Sequence 4, Appli
41	381	24.9	350	US-09-949-016-6526	Sequence 6526, Ap
42	381	24.9	377	US-09-949-016-11211	Sequence 11211, A
43	380	24.8	264	US-09-949-016-8464	Sequence 8464, Ap
44	333	21.8	350	US-08-852-481-2	Sequence 2, Appli
45	303.5	19.8	303	US-09-795-926-2	Sequence 2, Appli
46	303.5	19.8	303	US-10-364-774-2	Sequence 2, Appli
47	292.5	19.1	265	US-09-795-926-6	Sequence 6, Appli
48	292.5	19.1	265	US-10-364-774-6	Sequence 6, Appli
49	275.5	18.0	320	US-09-854-122-20	Sequence 20, Appli
50	265.5	17.4	324	US-09-854-122-18	Sequence 18, Appli
51	259.5	17.0	302	US-09-854-122-19	Sequence 19, Appli
52	254	16.6	214	US-09-795-926-16	Sequence 16, Appli
53	254	16.6	214	US-10-364-774-16	Sequence 16, Appli
54	187	12.2	186	US-09-795-926-12	Sequence 12, Appli
55	187	12.2	186	US-10-364-774-12	Sequence 12, Appli
56	187	12.2	331	US-09-854-122-17	Sequence 17, Appli
57	187	12.2	391	US-09-854-122-16	Sequence 16, Appli
58	176	11.5	148	US-09-795-926-8	Sequence 8, Appli
59	176	11.5	148	US-10-364-774-8	Sequence 8, Appli
60	150.5	9.8	106	US-10-094-749-2841	Sequence 2841, Ap
61	137.5	9.0	97	US-09-795-926-18	Sequence 18, Appli
62	137.5	9.0	97	US-10-364-774-18	Sequence 18, Appli
63	115	7.5	104	US-09-513-999C-8076	Sequence 8076, Ap
64	111	7.3	42	US-09-167-681-49	Sequence 49, Appli
65	106.5	7.0	325	US-10-126-279-25	Sequence 25, Appli
66	106.5	7.0	325	US-10-286-606-25	Sequence 25, Appli
67	106.5	7.0	326	US-10-891-383A-25	Sequence 25, Appli
68	104.5	6.8	148	US-09-795-926-10	Sequence 10, Appli
69	104.5	6.8	148	US-10-364-774-10	Sequence 10, Appli
70	104.5	6.8	1402	US-09-125-635-12	Sequence 12, Appli
71	101	6.6	42	US-09-167-681-41	Sequence 41, Appli
72	100.5	6.6	1420	US-09-125-635-4	Sequence 4, Appli
73	99	6.5	32	US-09-167-681-34	Sequence 34, Appli
74	99	6.5	32	US-09-167-681-42	Sequence 42, Appli
75	99	6.5	32	US-09-167-681-50	Sequence 50, Appli
76	97.5	6.4	907	US-09-949-016-10609	Sequence 10609, A
77	94	6.1	36	US-09-167-681-36	Sequence 36, Appli
78	94	6.1	37	US-09-167-681-44	Sequence 44, Appli
79	94	6.1	37	US-09-167-681-52	Sequence 52, Appli
80	94	6.1	42	US-09-167-681-33	Sequence 33, Appli
81	93.5	6.1	110	US-09-795-926-4	Sequence 4, Appli
82	93.5	6.1	110	US-10-364-774-4	Sequence 4, Appli
83	93.5	6.1	254	US-09-586-106D-69	Sequence 69, Appli
84	93.5	6.1	254	US-10-799-870-69	Sequence 69, Appli
85	92.5	6.0	883	US-09-949-016-6398	Sequence 6398, Ap
86	92.5	6.0	892	US-09-949-016-7255	Sequence 7255, Ap
87	88	5.8	121	US-09-270-767-41537	Sequence 41537, A
88	88	5.8	610	US-09-949-016-8271	Sequence 8271, Ap
89	88	5.8	638	US-09-949-016-8270	Sequence 8270, Ap
90	88	5.8	702	US-09-949-016-7288	Sequence 7288, Ap
91	88	5.8	752	US-08-244-189-2	Sequence 2, Appli
92	87.5	5.7	172	US-09-345-473E-12	Sequence 12, Appli
93	87.5	5.7	172	US-09-862-027-12	Sequence 12, Appli
94	87	5.7	154	US-08-946-329A-49	Sequence 49, Appli
95	87	5.7	154	US-09-562-914-49	Sequence 49, Appli
96	86.5	5.7	499	US-09-248-796A-14484	Sequence 14484, A
97	86	5.6	1402	US-09-445-333E-2	Sequence 2, Appli
98	85.5	5.6	241	US-10-094-749-2510	Sequence 2510, Ap
99	85.5	5.6	287	US-10-104-047-2485	Sequence 2485, Ap

```
100      84      5.5      254      2      US-09-586-106D-77      Sequence 77, Appl
ALIGNMENTS
RESULT 1
US-09-786-240-7
; Sequence 7, Application US/09786240
; Patent No. 6558935
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUGGLER, Karl J.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: AZIMZAI, Valda
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
; FILE REFERENCE: PF-0592 PCT
; CURRENT APPLICATION NUMBER: US/09/786,240
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/150,657; unassigned; 09/186,779; unassigned; 60/133,642
; PRIOR FILING DATE: 1998-09-10; 1998-09-10; 1998-11-04; 1998-11-04; 1999-05-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6558935 1784742CD1
US-09-786-240-7
```

```
Query Match      100.0%; Score 1530; DB 2; Length 284;
Best Local Similarity 100.0%; Pred. No. 3.9e-186;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1      MAESAETPSTGPESEKFFHGVRLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60
DB      1      MAESAETPSTGPESEKFFHGVRLPPFCRGKMEETANFVRPSDVWIVTPKSGTSL 60
QY      61      LQEVVYVVSQADPDEIGLMDIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
DB      61      LQEVVYVVSQADPDEIGLMDIDEQLPVLEYPQGLDIIKELTSPRLIKSHLPYRFLPSD 120
QY      121     LHNGDSKVIYARNPKDLVSVYQPHRSRLTMSYRGTFQECRRFMNDKLGYSWFHVVQ 180
DB      121     LHNGDSKVIYARNPKDLVSVYQPHRSRLTMSYRGTFQECRRFMNDKLGYSWFHVVQ 180
QY      181     EFWEHRMDSNVLFKYEDHMDRLVTMVQLARFLGVSCDKAQLAETHCHQLVDQCCNA 240
DB      181     EFWEHRMDSNVLFKYEDHMDRLVTMVQLARFLGVSCDKAQLAETHCHQLVDQCCNA 240
QY      241     EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
DB      241     EALPVGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTFDFYL 284
```

```
RESULT 2
US-09-854-122-21
; Sequence 21, Application US/09854122
; Patent No. 6841718
; GENERAL INFORMATION:
; APPLICANT: ALBERTE, RANDALL S.
; APPLICANT: SMITH, ROBERT
; TITLE OF INVENTION: TRANSGENETIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
; FILE REFERENCE: PHA-007.01
; CURRENT APPLICATION NUMBER: US/09/854,122
; CURRENT FILING DATE: 2001-09-10
```

```
; PRIOR APPLICATION NUMBER: 60/202,529
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 21
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-854-122-21
Query Match      30.8%; Score 472; DB 2; Length 295;
Best Local Similarity 35.8%; Pred. No. 3.5e-45;
Matches 100; Conservative 52; Mismatches 101; Indels 26; Gaps 4;
QY      24      GVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSLQEVVYVVSQADPDEIGLMDID 83
DB      17      GVPLIKYFAEALGPLQSFQARPDDLINTYTPKSGTTWVSQILDMYQGGDLEKCHRAPIF 76
QY      84      EQLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYARNPKDLV 139
DB      77      VRVPLEVNDGPEPSGLETAKDTPPRLIKSHLPALLPQTLDDQKVKVYVARNPKDVA 136
QY      140     VSYQPHRSRLTMSYRGTFQECRRFMNDKLGYSWFHVVQEFWEHRMDSNVLFKYEDM 199
DB      137     VSYVHFHRMEKAHPGPGTWDTSFLEKFMAGEVSYGWSYQHVQEWELSRTHPVLVLFYEDM 196
QY      200     HRDLVTWVEQLARFLGVSCDKAQLAETHC-----HQLVDQCCNA 242
DB      197     KENPKREIQKILEFVGRSLPEETMDPMVOHTSFKEMKKNPMNTYTTVPQELMDHSIS--- 253
QY      243     LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
DB      254     -PFMRKGMAGDWKTTFTVAQNERFDADYAEKMGCSLSF 291
RESULT 3
US-09-328-174A-3
; Sequence 3, Application US/09328174A
; Patent No. 6448003
; GENERAL INFORMATION:
; APPLICANT: Kurth, Janice
; APPLICANT: Guida, Marco
; TITLE OF INVENTION: Genotyping Human Phenol Sulfotransferase
; FILE REFERENCE: 4389-6 (formerly SEQ-16P)
; CURRENT APPLICATION NUMBER: US/09/328,174A
; CURRENT FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 09/328,174
; PRIOR FILING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 295
; TYPE: PRT
; ORGANISM: H. sapiens
US-09-328-174A-3
Query Match      29.8%; Score 456; DB 2; Length 295;
Best Local Similarity 33.7%; Pred. No. 2.3e-43;
Matches 94; Conservative 58; Mismatches 101; Indels 26; Gaps 4;
QY      24      GVRLLPPFCRGKMEETANFVRPSDVWIVTPKSGTSLQEVVYVVSQADPDEIGLMDID 83
DB      17      GVPLIKYFAEALGPLQSFQARPDDLINTYTPKSGTTWVSQILDMYQGGDLEKCHRAPIF 76
QY      84      EQLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYARNPKDLV 139
DB      77      MRVPLEFKVDPGIPSGMETLKNTPAPRLKTHLPALLPQTLDDQKVKVYVARNPKDVA 136
QY      140     VSYQPHRSRLTMSYRGTFQECRRFMNDKLGYSWFHVVQEFWEHRMDSNVLFKYEDM 199
DB      137     VSYVHFYHMAKVYPHPGPTWESFLEKFMAGEVSYGWSYQHVQEWELSRTHPVLVLFYEDM 196
```



```
Qy 200 HRDLVTMVEQLARFLGVSCDKAQLALTEHC-----HQLVDQCCNABA 242
; SEQ ID NO 2
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-792-695A-2

Db 197 KENPKREIQKILEFVGRSLPEETVDLMVTEHTSFKEKMKNPMTNTYTTVRREFMDHSIS--- 253

Qy 243 LP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
; SEQ ID NO 243
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-792-695A-2

Db 254 -PFRKKGAGDWKTTFTVAQNERFDADYAEKMGCSLSF 291

RESULT 4
US-09-949-016-6018
; Sequence 6018, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6018
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6018

Query Match 29.3%; Score 449; DB 2; Length 296;
Best Local Similarity 33.1%; Pred. No. 1.5e-42;
Matches 94; Conservative 61; Mismatches 105; Indels 24; Gaps 4;

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6018
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6018

Qy 18 KYFEFHGVRLPPFCRGKMEETANFVRPSDVMIYTPKSGTSLQEVVYLVVSQGADPDEI 77
; SEQ ID NO 18
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-792-695A-2

Db 12 KLKEVEGTLQPATVDNWSQIQSFQKAPDDLLICTYPKAGITWIEIVDMIEQNGDVEKC 71

Qy 78 GLMNIDEQLPVLEY---POP-GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMAR 133
; SEQ ID NO 78
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-792-695A-2

Db 72 QRATIIQHRHPPIEWARPQPSGVEKAKAMPSPRILKTHLSTQLLPPSPWENCKFLYVAR 131

Qy 134 NPKDLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSGFHFVQEFWEHRMDSNVLF 193
; SEQ ID NO 134
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-792-695A-2

Db 132 NAKDCMVSYYHFORNNHMLPDGWTWEEYFETTINGVVMGWSFDFHVGKGMWKORHQLIF 191

Qy 194 LKYEDMHRDLVTMVEQLARFLGVSCDKAQLALTEH-----CHQLVDQ 236
; SEQ ID NO 194
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-792-695A-2

Db 192 LFYEDIKRDPKHEIRKVMQFNGKKVDETVLDKIYQETSFKEKMKNPMTNRTSVSKSLDQ 251

Qy 237 CCNAEALPVGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
; SEQ ID NO 237
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-792-695A-2

Db 252 SSSSF---MRKGTVDGDKNHFTVAQNERFDEIYRRKMEGTSLNF 292

RESULT 5
US-09-792-695A-2
; Sequence 2, Application US/09792695A
; Patent No. 7026163
; GENERAL INFORMATION:
; APPLICANT: Freimuth, Robert R.
; APPLICANT: Weinshilboum, Richard M.
; APPLICANT: Wieden, Eric D.
; TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS
; FILE REFERENCE: 07039-273001
; CURRENT APPLICATION NUMBER: US/09/792,695A
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 2
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-792-695A-2

Query Match 29.3%; Score 449; DB 3; Length 296;
Best Local Similarity 33.1%; Pred. No. 1.5e-42;
Matches 94; Conservative 61; Mismatches 105; Indels 24; Gaps 4;

Qy 18 KYFEFHGVRLPPFCRGKMEETANFVRPSDVMIYTPKSGTSLQEVVYLVVSQGADPDEI 77
; SEQ ID NO 18
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-792-695A-2

Db 12 KLKEVEGTLQPATVDNWSQIQSFQKAPDDLLICTYPKAGITWIEIVDMIEQNGDVEKC 71

Qy 78 GLMNIDEQLPVLEY---POP-GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMAR 133
; SEQ ID NO 78
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-792-695A-2

Db 72 QRATIIQHRHPPIEWARPQPSGVEKAKAMPSPRILKTHLSTQLLPPSPWENCKFLYVAR 131

Qy 134 NPKDLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSGFHFVQEFWEHRMDSNVLF 193
; SEQ ID NO 134
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-792-695A-2

Db 132 NAKDCMVSYYHFORNNHMLPDGWTWEEYFETTINGVVMGWSFDFHVGKGMWKORHQLIF 191

Qy 194 LKYEDMHRDLVTMVEQLARFLGVSCDKAQLALTEH-----CHQLVDQ 236
; SEQ ID NO 194
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-792-695A-2

Db 192 LFYEDIKRDPKHEIRKVMQFNGKKVDETVLDKIYQETSFKEKMKNPMTNRTSVSKSLDQ 251

Qy 237 CCNAEALPVGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
; SEQ ID NO 237
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-792-695A-2

Db 252 SSSSF---MRKGTVDGDKNHFTVAQNERFDEIYRRKMEGTSLNF 292

RESULT 6
US-09-949-016-10328
; Sequence 10328, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10328
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10328

Query Match 28.4%; Score 435; DB 2; Length 309;
Best Local Similarity 32.7%; Pred. No. 6.3e-41;
Matches 96; Conservative 59; Mismatches 121; Indels 18; Gaps 3;

Qy 5 EAETPSTPGFESKYFFHGVRLPPFCRGKMEETANFVRPSDVMIYTPKSGTSLQEV 64
; SEQ ID NO 5
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-949-016-10328

Db 12 QCTTMSELDDYKEFEEVHGILMYKDFVKYVNDVNEAFQARPPDDIVATYIPKSGTWWSEI 71

Qy 65 VYLVVSQGADPDEIGLMNIDEQLPVLEYPOP---GLDIIKELTSPRLIKSHLPYRFLPSD 120
; SEQ ID NO 65
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-949-016-10328

Db 72 VIMIYKGGDVEKCKEDVIFNRIPLECKENMLMGVQLQDMMNSPRIVKTHLPPELLPAS 131

Qy 121 LHNGDSKVIYMARNPDLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSGFHFVQ 180
; SEQ ID NO 121
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-949-016-10328

Db 132 FWEKDKIILYCRNAKDVAVSYFFFLMWAGHPNPGSFPEFVEKFMQGVPGYSWKYHKV 191

Qy 181 EFWHRMDSNVLFUKYEDMHRDLVTMVEQLARFLGVSCDKAQLALTEHCH----- 231
```

```

192 SWEKGKSPRVLFYEDLDKEVILKIHFLERKPSBELVDRIIHTTSFOEMKNKPS 251
193
232 ----QLVDDCCNABALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
194
252 TNYTTLPEIMNQKLSPFMRKGITGDKWKNHFTVALNEKFDKHVEQOMKESTLKF 305
195
RESULT 7
US-08-325-562-2
; Sequence 2, Application US/08325562
; Patent No. 5714594
; GENERAL INFORMATION:
; APPLICANT: Weinshilboun, Richard M.
; APPLICANT: Aksoy, Ibrahim A.
; APPLICANT: Wood, Thomas C.
; TITLE OF INVENTION: CDNA CLONING AND EXPRESSION OF HUMAN
; TITLE OF INVENTION: LIVER ESTROGEN SULFOTRANSFERASE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg & Woessner P.A.
; STREET: 3500 IDS Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/325.562
; FILING DATE: 18-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mueiting, Ann M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 150.131US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-325-562-2

Query Match 28.4%; Score 434; DB 1; Length 294;
Best Local Similarity 33.3%; Pred. No. 7.6e-41;
Matches 95; Conservative 56; Mismatches 116; Indels 18; Gaps 3;

Qy 14 EFESKYFEFHGVRLPPFCRGKMEIANFPVRPSDVMTVTPKSGTSLQLQEVVYLVSGAD 73
Db 6 DYEFKEFEVHGILMYKDFVKYNDVNEAFQARPDLDVIATYPSGTTWVSEIVYMYKEGD 65
Qy 74 PDEIGLNMIDQLPVLEYQPQ-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVI 129
Db 66 VEKCKEDVFNRIFFLECRKENLMGVKQDDEMNSPRIVKTHLPPELLPASFWEKDCCKII 125
Qy 130 YMAENPKDLVSYVQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVEQFWEHRMDS 189
Db 126 YLCRNAKDVAVSFYFFFLMVAGHPNPGSPFPEFVEKFMQGVPPYGSWKYKHWKMGKSP 185
Qy 190 NVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLTEALTEHCH-----QLVDQ 236
Db 186 RVLFYEDLKEDIRKEVILKIHFLERKPSBELVDRIIHTTSFOEMKNKPSNTYTLTLPDE 245
Qy 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 246 IMNQKLSPFMRKGITGDKWKNHFTVALNEKFDKHVEQOMKESTLKF 290

RESULT 9
US-09-150-133-13
; Sequence 13, Application US/09150133B
; Patent No. 6060295
```

```

192 SWEKGKSPRVLFYEDLDKEVILKIHFLERKPSBELVDRIIHTTSFOEMKNKPS 251
193
232 ----QLVDDCCNABALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
194
252 TNYTTLPEIMNQKLSPFMRKGITGDKWKNHFTVALNEKFDKHVEQOMKESTLKF 305
195
RESULT 8
US-08-437-795-2
; Sequence 2, Application US/08437795
; Patent No. 5744355
; GENERAL INFORMATION:
; APPLICANT: Weinshilboun, Richard M.
; APPLICANT: Aksoy, Ibrahim A.
; APPLICANT: Wood, Thomas C.
; TITLE OF INVENTION: CDNA CLONING AND EXPRESSION OF HUMAN
; TITLE OF INVENTION: LIVER ESTROGEN SULFOTRANSFERASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg & Woessner P.A.
; STREET: 3500 IDS Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/437,795
; FILING DATE: 09-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mueiting, Ann M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 150.137US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-437-795-2

Query Match 28.4%; Score 434; DB 1; Length 294;
Best Local Similarity 33.3%; Pred. No. 7.6e-41;
Matches 95; Conservative 56; Mismatches 116; Indels 18; Gaps 3;

Qy 14 EFESKYFEFHGVRLPPFCRGKMEIANFPVRPSDVMTVTPKSGTSLQLQEVVYLVSGAD 73
Db 6 DYEFKEFEVHGILMYKDFVKYNDVNEAFQARPDLDVIATYPSGTTWVSEIVYMYKEGD 65
Qy 74 PDEIGLNMIDQLPVLEYQPQ-----GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVI 129
Db 66 VEKCKEDVFNRIFFLECRKENLMGVKQDDEMNSPRIVKTHLPPELLPASFWEKDCCKII 125
Qy 130 YMAENPKDLVSYVQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVEQFWEHRMDS 189
Db 126 YLCRNAKDVAVSFYFFFLMVAGHPNPGSPFPEFVEKFMQGVPPYGSWKYKHWKMGKSP 185
Qy 190 NVLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLTEALTEHCH-----QLVDQ 236
Db 186 RVLFYEDLKEDIRKEVILKIHFLERKPSBELVDRIIHTTSFOEMKNKPSNTYTLTLPDE 245
Qy 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 246 IMNQKLSPFMRKGITGDKWKNHFTVALNEKFDKHVEQOMKESTLKF 290

RESULT 9
US-09-150-133-13
; Sequence 13, Application US/09150133B
; Patent No. 6060295
```

```
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES, NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 5820.504
; CURRENT APPLICATION NUMBER: US/09/150,133B
; CURRENT FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 5.1 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-150-133-13

Query Match      28.1%; Score 430; DB 2; Length 295;
Best Local Similarity 33.0%; Pred. No. 2.2e-40;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYFHFHGVRLPPFCRCGMKEIANFPVRPSDVMIIVTPKSGTSLQLQEVVYLVSQGAD 73
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 7 EYEVFGFGRGLVMDKRTKYWEDVEMFLARPDDLVIATYPKSGTTWISVVYMIYKEGD 66

QY 74 PDEIGLMNIDQLPVLEYQPQ-----GLDIIKELTSPLRIKSHLPYRFLPDLHNGDSKVI 129
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 67 VEKCKEDAFNRIPLYLECRNEDLINGIKQLKEKESPRIVKTHLPKVLPPASFWEKNCXKI 126

QY 130 YMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDS 189
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 127 YLCRNAKDVAVSYYYFLMLMITSYNPNKSFSEFEKFMQGVPPYGSWDHVKAWEKSKNS 186

QY 190 NVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCH-----QLVDQ 236
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 187 RVLFMFYEDMKEDIRREVVKLIEFLERKPSAELVDRIIQTSTFQEMKKNPSTNYTMPEE 246

QY 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 247 MNQKVSPFMRKGIIGDKWNHFPALRERFDEHYKQMKDCTVKF 291

; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES, AND METHODS OF USE THEREOF
; FILE REFERENCE: 5820.546
; CURRENT APPLICATION NUMBER: US/09/374,493
; CURRENT FILING DATE: 1999-08-13
; EARLIER APPLICATION NUMBER: 09/150,133
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 60/072,994
; EARLIER FILING DATE: 1998-04-29
; EARLIER APPLICATION NUMBER: PCT/US99/16750
; EARLIER FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-374-493-13

Query Match      28.1%; Score 430; DB 2; Length 295;
Best Local Similarity 33.0%; Pred. No. 2.2e-40;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYFHFHGVRLPPFCRCGMKEIANFPVRPSDVMIIVTPKSGTSLQLQEVVYLVSQGAD 73
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 7 EYEVFGFGRGLVMDKRTKYWEDVEMFLARPDDLVIATYPKSGTTWISVVYMIYKEGD 66

QY 74 PDEIGLMNIDQLPVLEYQPQ-----GLDIIKELTSPLRIKSHLPYRFLPDLHNGDSKVI 129
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 67 VEKCKEDAFNRIPLYLECRNEDLINGIKQLKEKESPRIVKTHLPKVLPPASFWEKNCXKI 126

QY 130 YMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDS 189
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 127 YLCRNAKDVAVSYYYFLMLMITSYNPNKSFSEFEKFMQGVPPYGSWDHVKAWEKSKNS 186

QY 190 NVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCH-----QLVDQ 236
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 187 RVLFMFYEDMKEDIRREVVKLIEFLERKPSAELVDRIIQTSTFQEMKKNPSTNYTMPEE 246

QY 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 247 MNQKVSPFMRKGIIGDKWNHFPALRERFDEHYKQMKDCTVKF 291

; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES, AND METHODS OF USE THEREOF
; FILE REFERENCE: 5820.495
; CURRENT APPLICATION NUMBER: US/09/150,141B
; CURRENT FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 5.1 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-150-141-13

Query Match      28.1%; Score 430; DB 2; Length 295;
Best Local Similarity 33.0%; Pred. No. 2.2e-40;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYFHFHGVRLPPFCRCGMKEIANFPVRPSDVMIIVTPKSGTSLQLQEVVYLVSQGAD 73
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 7 EYEVFGFGRGLVMDKRTKYWEDVEMFLARPDDLVIATYPKSGTTWISVVYMIYKEGD 66

QY 74 PDEIGLMNIDQLPVLEYQPQ-----GLDIIKELTSPLRIKSHLPYRFLPDLHNGDSKVI 129
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
```

```
Db 67 VEKCKEDAFNRIPLYLECRNEDLINGIKQLKEKESPRIVKTHLPKVLPPASFWEKNCXKI 126
QY 130 YMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDS 189
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 127 YLCRNAKDVAVSYYYFLMLMITSYNPNKSFSEFEKFMQGVPPYGSWDHVKAWEKSKNS 186
QY 190 NVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCH-----QLVDQ 236
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 187 RVLFMFYEDMKEDIRREVVKLIEFLERKPSAELVDRIIQTSTFQEMKKNPSTNYTMPEE 246
QY 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 247 MNQKVSPFMRKGIIGDKWNHFPALRERFDEHYKQMKDCTVKF 291

RESULT 11
US-09-374-493-13
; Sequence 13, Application US/09374493
; Patent No. 6204016
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5820-546
; CURRENT APPLICATION NUMBER: US/09/374,493
; CURRENT FILING DATE: 1999-08-13
; EARLIER APPLICATION NUMBER: 09/150,133
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 60/072,994
; EARLIER FILING DATE: 1998-04-29
; EARLIER APPLICATION NUMBER: PCT/US99/16750
; EARLIER FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-374-493-13

Query Match      28.1%; Score 430; DB 2; Length 295;
Best Local Similarity 33.0%; Pred. No. 2.2e-40;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYFHFHGVRLPPFCRCGMKEIANFPVRPSDVMIIVTPKSGTSLQLQEVVYLVSQGAD 73
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 7 EYEVFGFGRGLVMDKRTKYWEDVEMFLARPDDLVIATYPKSGTTWISVVYMIYKEGD 66

QY 74 PDEIGLMNIDQLPVLEYQPQ-----GLDIIKELTSPLRIKSHLPYRFLPDLHNGDSKVI 129
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 67 VEKCKEDAFNRIPLYLECRNEDLINGIKQLKEKESPRIVKTHLPKVLPPASFWEKNCXKI 126

QY 130 YMARNPDKLVVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFEHVOEFWEHRMDS 189
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 127 YLCRNAKDVAVSYYYFLMLMITSYNPNKSFSEFEKFMQGVPPYGSWDHVKAWEKSKNS 186

QY 190 NVFLFKYEDMHRDLVTWVEQLARFLGVSCDKAQLEALTEHCH-----QLVDQ 236
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 187 RVLFMFYEDMKEDIRREVVKLIEFLERKPSAELVDRIIQTSTFQEMKKNPSTNYTMPEE 246
QY 237 CCNAEALP-VGRGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 247 MNQKVSPFMRKGIIGDKWNHFPALRERFDEHYKQMKDCTVKF 291

RESULT 12
US-09-374-824-13
; Sequence 13, Application US/09374824
; Patent No. 6207414
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
```

[illegible]

```

RESULT 15
US-10-411-976-13
; Sequence 13, Application US/10411976
; Patent No. 6713283
; GENERAL INFORMATION:
; APPLICANT: Moore, Kevin L.
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES
; FILE REFERENCE: 5864.027
; CURRENT APPLICATION NUMBER: US/10/411.976
; PRIOR FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 09/785,343
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: estrogen sulfotransferase
US-10-411-976-13

Query Match      28.1%; Score 430; DB 2; Length 295;
Best Local Similarity 33.0%; Pred.No.2.2e-40;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYEFHGVRLLPPFCRGKMBEIANFPVRPSDWIIVTPKSGTSLLOEYVILVSGAD 73
DB 7 EYEVFGFGRGLMDKRFYKRWEDVFLARPDDDLVIATYTPKSGTTWISVAVYMIYKEGD 66
QY 74 PDEIGLMNIDBOLPVLRYQPQ---GLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVI 129
DB 67 VEKCKEDAIENRPLYSCRNEDLINGIKQKEKESPRIVKTHLPKPLPASFWKKNCKMI 126
QY 130 YMARNPNDLVVSYQFHSRLTMSYRGTFQBFCCRFMNDKLGYSWFEHQEFWEHRMDS 189
DB 127 YLCENAKDAVSYFYFLMITSYPNPKSFSEFEKFMQGVPGYGSWDYDHVKANWEKSNS 186
QY 190 NVLFLKYEDMHRDLVTVWQELARPLGVSCDKAQLEALTEHCH-----QLVDQ 236
DB 187 RVLFMFYEDMKEDIRREVVWKLIEFLERKPSAELVDRIIOHTSFQEMKNPSTNYTMPEE 246
QY 237 CCNAEALP-VGRGVLGWLKDIFTVSMNKEFDLVYKQMGKCDLTF 280
DB 247 MMNQKVSPPMRKGIIGDQKNHFPALRERFDEHYKQMKDCTVKF 291

RESULT 16
US-10-201-525-13
; Sequence 13, Application US/10201525
; Patent No. 7029890
; GENERAL INFORMATION:
; APPLICANT: Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5827.005
; CURRENT APPLICATION NUMBER: US/10/201.525
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/785,343
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: PCT/US99/16750
; PRIOR FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-201-525-13

Query Match      28.1%; Score 430; DB 3; Length 295;
Best Local Similarity 33.0%; Pred.No.2.2e-40;
Matches 94; Conservative 57; Mismatches 116; Indels 18; Gaps 3;

QY 14 EFESKYEFHGVRLLPPFCRGKMBEIANFPVRPSDWIIVTPKSGTSLLOEYVILVSGAD 73

```

```
RESULT 18
US-10-199-334-5
; Sequence 5, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT FILING DATE: 2002-07-22
; CURRENT APPLICATION NUMBER: US/10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-5

Query Match 27.6%; Score 422; DB 2; Length 304;
Best Local Similarity 32.3%; Pred. No. 1.9e-39;
Matches 97; Conservative 59; Mismatches 124; Indels 20; Gaps 5;

QY 1 MAESAETPTSGEPE-SKYFEFHGVRLLPPFCRGKMEBIANFPVRPSDVWIVTPKSGTS 59
Db 1 MAKIEKNAPTMEKKPELFNMEVDGVPVTLILSKWEKVCNFOAKPDDLILATPKSGTT 60
QY 60 LLOEVVYLVISQAGADPEIGLNMIDQLPVLE--YP---QPGLDIKELTSPRLIKSHLPY 114
Db 61 WMEHILDMILNDGDVEKCKRAQTLDRHAFLEKFPHEKPDLEFVLEMSSPQLIKTHLPS 120
QY 115 RFLPSDLHNGDSKVIYMARNPDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYS 174
Db 121 HLIPPSIWKENCKIVYVARNPKDCLVSYVHFRMASFMPDPQNLEEFYEFKFMGKVGGS 180
QY 175 WFEHVQEFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLALETCHQLV 234
Db 181 WPDVKVGWAAKDMHRILYLFYEDIKKDPKREIKILKFLKEDISEELINKIYHTSPDV 240
QY 235 ---DQCCNAEALPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 MKQNPMTNYYTLPTSIMDHSISPFMRKMGPDGKNYFTVAQNEEFKDYQKQKMGSTLTF 300

RESULT 20
US-10-199-330-5
; Sequence 5, Application US/10199330
; Patent No. 6967094
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-5

Query Match 27.6%; Score 422; DB 2; Length 304;
Best Local Similarity 32.3%; Pred. No. 1.9e-39;
Matches 97; Conservative 59; Mismatches 124; Indels 20; Gaps 5;

QY 1 MAESAETPTSGEPE-SKYFEFHGVRLLPPFCRGKMEBIANFPVRPSDVWIVTPKSGTS 59
Db 1 MAKIEKNAPTMEKKPELFNMEVDGVPVTLILSKWEKVCNFOAKPDDLILATPKSGTT 60
QY 60 LLOEVVYLVISQAGADPEIGLNMIDQLPVLE--YP---QPGLDIKELTSPRLIKSHLPY 114
Db 61 WMEHILDMILNDGDVEKCKRAQTLDRHAFLEKFPHEKPDLEFVLEMSSPQLIKTHLPS 120
QY 115 RFLPSDLHNGDSKVIYMARNPDLVVSYYQFHRSLRTMSYRGTFQFCRRFMNDKLGYS 174
Db 121 HLIPPSIWKENCKIVYVARNPKDCLVSYVHFRMASFMPDPQNLEEFYEFKFMGKVGGS 180
QY 175 WFEHVQEFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLALETCHQLV 234
Db 181 WPDVKVGWAAKDMHRILYLFYEDIKKDPKREIKILKFLKEDISEELINKIYHTSPDV 240
QY 235 ---DQCCNAEALPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 MKQNPMTNYYTLPTSIMDHSISPFMRKMGPDGKNYFTVAQNEEFKDYQKQKMGSTLTF 300

RESULT 19
US-10-199-329-5
; Sequence 5, Application US/10199329
; Patent No. 6953681
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
```

```
QY 115 RFLPSDLHNGDSKVIYMARNPDKDLVVSYYQFHRSLRTMSYRGTFQFCRRFMDNKLGVGS 174
Db 121 HLIPIPSIWKENCKIVVVAENPKDCLVSYHYHFRMASFMPDPQNLBEFYEKFMGKVVVGS 180
QY 175 WFEHVQEFWEHRMDSNVLKYEYDHRDLVTVVVEQLARFLGVSCDKAQLEALTEHCHQLV 234
Db 181 WFDHVGWGAADKMRILYLFYEDIKDKPKRIEKLKLEKDISEILNKIIYHTSFV 240
QY 235 ---DQCNAEALPVG-----RGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
Db 241 MKQNPWTNVTTLPTSIMDHSISPFMRKMGPGDKNVTVAQNEEFDKYQKMGAGSTLTF 300

RESULT 21
US-09-609-816-8
; Sequence 8, Application US/09609816
; Patent No. 6436684
; GENERAL INFORMATION:
; APPLICANT: Woodage, Trevor
; APPLICANT: Wei, Minh Hui
; APPLICANT: Kodira, Chinappa
; APPLICANT: Beasley, Ellen
; APPLICANT: DiFrancesco, Valentina
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669PCT
; CURRENT APPLICATION NUMBER: US/09/609,816
; CURRENT FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: HUMAN
US-09-609-816-8

Query Match 27.4%; Score 419.5; DB 2; Length 283;
Best Local Similarity 32.6%; Pred. No. 3.3e-39;
Matches 91; Conservative 57; Mismatches 112; Indels 19; Gaps 4;

QY 21 EFHGVLPPFCRGKMEETANFPVRPSDVWIVTYPKSGTSLLOEVYLVLSQGADPDPEIGLM 80
Db 1 EVNGILMSKLSMDNWKIWNFOAKPDDLLIATYAKAGTTWTQEI VDMIQNDGVDQKQORA 60
QY 81 NIDEOLPVLEY--POP---GLDIIKELTSPRLIKSHLPVRLPSPDLHNGDSKVIYMARNP 135
Db 61 NYDRHPFIEWTLPSPLNSGLDLANKMSPRTLKTHLPVHMLPPSFWEKNSKIIIVARNA 120
QY 136 KDLVVSYYQFHRSLRTMSYRGTFQFCRRFMDNKLGVGSWFHVEHQRMDSNVLFLK 195
Db 121 KDLVSYYYFSSRMNKMPLDPGTGLGEYIEQFKAGKVLGWSYDHWKGVMDVKDQHRILYLF 180
QY 196 YEDMHRDLVTVVVEQLARFLGVSCDKAQLEALTEHCHQLV---DQCNAEALP----- 244
Db 181 YEDMKDKPKRIEKKIAKLEKDISEVLNKIIYHTSFVDMKENPNMANYTTLPSSIMDHSI 240
QY 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
Db 241 SPFMRKMGPGDKNVTVAQSEDFEDYRRKMGAGSNITF 279

RESULT 22
US-09-609-816-9
; Sequence 9, Application US/09609816
; Patent No. 6436684
```

```
; GENERAL INFORMATION:
; APPLICANT: Woodage, Trevor
; APPLICANT: Wei, Minh Hui
; APPLICANT: Kodira, Chinappa
; APPLICANT: Beasley, Ellen
; APPLICANT: DiFrancesco, Valentina
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669PCT
; CURRENT APPLICATION NUMBER: US/09/609,816
; CURRENT FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 283
; TYPE: PRT
; ORGANISM: HUMAN
US-09-609-816-9

Query Match 27.4%; Score 419.5; DB 2; Length 283;
Best Local Similarity 32.6%; Pred. No. 3.3e-39;
Matches 91; Conservative 57; Mismatches 112; Indels 19; Gaps 4;

QY 21 EFHGVLPPFCRGKMEETANFPVRPSDVWIVTYPKSGTSLLOEVYLVLSQGADPDPEIGLM 80
Db 1 EVNGILMSKLSMDNWKIWNFOAKPDDLLIATYAKAGTTWTQEI VDMIQNDGVDQKQORA 60
QY 81 NIDEOLPVLEY--POP---GLDIIKELTSPRLIKSHLPVRLPSPDLHNGDSKVIYMARNP 135
Db 61 NYDRHPFIEWTLPSPLNSGLDLANKMSPRTLKTHLPVHMLPPSFWEKNSKIIIVARNA 120
QY 136 KDLVVSYYQFHRSLRTMSYRGTFQFCRRFMDNKLGVGSWFHVEHQRMDSNVLFLK 195
Db 121 KDLVSYYYFSSRMNKMPLDPGTGLGEYIEQFKAGKVLGWSYDHWKGVMDVKDQHRILYLF 180
QY 196 YEDMHRDLVTVVVEQLARFLGVSCDKAQLEALTEHCHQLV---DQCNAEALP----- 244
Db 181 YEDMKDKPKRIEKKIAKLEKDISEVLNKIIYHTSFVDMKENPNMANYTTLPSSIMDHSI 240
QY 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
Db 241 SPFMRKMGPGDKNVTVAQSEDFEDYRRKMGAGSNITF 279

RESULT 23
US-10-199-334-8
; Sequence 8, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: Woodage, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000869DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
```





```
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-9

Query Match      27.4%; Score 419.5; DB 2; Length 283;
Best Local Similarity 32.6%; Pred. No. 3.3e-39;
Matches 91; Conservative 57; Mismatches 112; Indels 19; Gaps 4;

Qy 21 EFHGVRLPPFCRGKMEETIANFVRPSDVWIVTYPKSGTSLQEVVYLVSQGADPDEIGLM 80
Db 1 EVNGILMSKLSMDNDWKIWNFOAKPDDLLIATYAKAGTTWTQEIIVDMIQNDGDVQKQORA 60

Qy 81 NIDEQLPVLEY--POP---GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
Db 61 NYDRHPFIEWTLPSPLNSGLDLANKMPSRTLKTHLPVHMLPPSFWKENSIIYVARNA 120

Qy 136 KDLVVSYQFHRSLRTMSYRGTFQFCRRFNMNDKLGYSWFHFVQEFWEHRMDSNVLFK 195
Db 121 KDLVSYYYFSRMNMKMLPDGTLGEYIEQFKAGKVLGMSWDYHVKGWMDVQDQRILYLF 180

Qy 196 YEDMRDLVTWVEQLARFLGVSCDKAQLEALTECHQLV---DQCCNAEALP----- 244
Db 181 YEDMKEDPKREIKKIAKLEKDISEEVNLKIIYHTSPDVMKENPMANYTTLPSSIMDHSI 240

Qy 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
Db 241 SPFMRKMGPGDWKNYFTVAQSEDFEDYRRKWAGSNITF 279

RESULT 28
US-10-199-330-9
; Sequence 9, Application US/10199330
; Patent No. 6967094
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-9

Query Match      27.4%; Score 419.5; DB 2; Length 283;
Best Local Similarity 32.6%; Pred. No. 3.3e-39;
Matches 91; Conservative 57; Mismatches 112; Indels 19; Gaps 4;

Qy 21 EFHGVRLPPFCRGKMEETIANFVRPSDVWIVTYPKSGTSLQEVVYLVSQGADPDEIGLM 80
Db 1 EVNGILMSKLSMDNDWKIWNFOAKPDDLLIATYAKAGTTWTQEIIVDMIQNDGDVQKQORA 60

Qy 81 NIDEQLPVLEY--POP---GLDIIKELTSPLRIKSHLPYRFLPSDLHNGDSKVIYMARNP 135
Db 61 NYDRHPFIEWTLPSPLNSGLDLANKMPSRTLKTHLPVHMLPPSFWKENSIIYVARNA 120

Qy 136 KDLVVSYQFHRSLRTMSYRGTFQFCRRFNMNDKLGYSWFHFVQEFWEHRMDSNVLFK 195
Db 121 KDLVSYYYFSRMNMKMLPDGTLGEYIEQFKAGKVLGMSWDYHVKGWMDVQDQRILYLF 180

Qy 196 YEDMRDLVTWVEQLARFLGVSCDKAQLEALTECHQLV---DQCCNAEALP----- 244
Db 181 YEDMKEDPKREIKKIAKLEKDISEEVNLKIIYHTSPDVMKENPMANYTTLPSSIMDHSI 240

Qy 245 ---VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
Db 241 SPFMRKMGPGDWKNYFTVAQSEDFEDYRRKWAGSNITF 279

RESULT 27
US-10-199-330-8
; Sequence 8, Application US/10199330
; Patent No. 6967094
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-8

Query Match      27.4%; Score 419.5; DB 2; Length 283;
```



```
Db 1 EVNGILSKMSENWDKIWNFOAKPDDLLIATYAKAGTTWTQEIYDMLQNDGDVQKQORA 60
Qy 81 NIDBOLPVLVYQPQ-----GLDIIKELTSPRLIKSHLPYRFLPSDLNNGDSKVYIMARNP 135
Db 61 NTDYRHPPIEWTLPPPLNSGLDLANKMPSRPTLTKTHLPVQMLPPSPFWKENSQIIYVARNA 120
Qy 136 KDLVVSYQYFHRSLRTMSYRGTFOFCRRFNMNDKLGYSWFEHVOEFWEHRMDSNVLPFLK 195
Db 121 KDLVSYYYFHRMKNMLPDGTLGEYITTFKAGVLMGWSYDHWKGVMDVVKDKHRIYLF 180
Qy 196 YEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLV---DOCCNAEALP----- 244
Db 181 YEDMKDKPKREIKKIVKLEKDISEEVLNKIIHTSPDMQKPNMANYTTLPPSSIMDHSI 240
Qy 245 ---VGRGVLGWKOIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 SPFMRKMGPDGKNYFTVAQSEDFEDYRKVMAGSTITF 279

RESULT 32
US-10-199-330-10
; Sequence 10, Application US/10199330
; Patent No. 6967094
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-10

Query Match 26.7%; Score 408.5; DB 2; Length 283;
Best Local Similarity 31.5%; Pred. No. 5.9e-38;
Matches 88; Conservative 57; Mismatches 115; Indels 19; Gaps 3;

Qy 21 EFHGVRLPPFCRGKMEETANFPVRPSDVWIVTPYKSGTSLQGVVYLVSQGADPDEIGLM 80
Db 1 EVNGILSKMSENWDKIWNFOAKPDDLLIATYAKAGTTWTQEIYDMLQNDGDVQKQORA 60
Qy 81 NIDBOLPVLVYQPQ-----GLDIIKELTSPRLIKSHLPYRFLPSDLNNGDSKVYIMARNP 135
Db 61 NTDYRHPPIEWTLPPPLNSGLDLANKMPSRPTLTKTHLPVQMLPPSPFWKENSQIIYVARNA 120
Qy 136 KDLVVSYQYFHRSLRTMSYRGTFOFCRRFNMNDKLGYSWFEHVOEFWEHRMDSNVLPFLK 195
Db 121 KDLVSYYYFHRMKNMLPDGTLGEYITTFKAGVLMGWSYDHWKGVMDVVKDKHRIYLF 180
Qy 196 YEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLV---DOCCNAEALP----- 244
Db 181 YEDMKDKPKREIKKIVKLEKDISEEVLNKIIHTSPDMQKPNMANYTTLPPSSIMDHSI 240
Qy 245 ---VGRGVLGWKOIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 SPFMRKMGPDGKNYFTVAQSEDFEDYRKVMAGSTITF 279

RESULT 33
```

```
US-09-609-816-6
; Sequence 6, Application US/09609816
; Patent No. 6436684
; GENERAL INFORMATION:
; APPLICANT: Woodage, Trevor
; APPLICANT: Wei, Minh Hui
; APPLICANT: Kodira, Chinappa
; APPLICANT: Beasley, Ellen
; APPLICANT: DiFrancesco, Valentina
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669PCT
; CURRENT APPLICATION NUMBER: US/09/609,816
; CURRENT FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 304
; TYPE: PRT
; ORGANISM: HUMAN
US-09-609-816-6

Query Match 26.6%; Score 407; DB 2; Length 304;
Best Local Similarity 31.3%; Pred. No. 9.8e-38;
Matches 94; Conservative 61; Mismatches 125; Indels 20; Gaps 5;

Qy 1 MAESEAEPTPGGFE-SKYFEPHGVRLPPFCRGKMBEIANFPVRPSDVWIVTPYKSGTS 59
Db 1 MAKTEKNAPTMEKKPELNFMEVDGVPVTLILSKWEKVCNFOAKPDDLLIATPKSGTT 60
Qy 60 LLQEVVYLVSQGADPDEIGLMNIDEQLPVLE--VP---QPGLDIIKELTSPRLIKSHLPY 114
Db 61 WMHEILDMLNDGDVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSPLIKTHLPS 120
Qy 115 RFLPSDLHNGDSKVYIMARNPKOLVSVYQYFHRSLRTMSYRGTFOFCRRFNMNDKLGYS 174
Db 121 HLPPSIWKENCKIVYVARNPCKCLSVYHFRHMASFMPDPQNLLEEFYEKFMGKVGG 180
Qy 175 WFEHVQEFWEHRMDSNVLFKLYEDMDHRLDVTWVQLARFLGVSCDKAQLAEALTEHCHQLV 234
Db 181 WFDHVGWAAKQDTHRIYLYFYEDIKKNPKRHEIHKVLEFLEKTLSGDVINKIVHHTSFDV 240
Qy 235 ---DOCCNAEALP-----VGRGVLWKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
Db 241 MKONPMANTAVPAHIFNHSISKFMKMGPMGDMKNHFTVAMNENFDKHYEKKMAGSTLNF 300

RESULT 34
US-10-199-334-6
; Sequence 6, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
```



```
; APPLICANT: Wei, Minh Hui
; APPLICANT: Kodira, Chinappa
; APPLICANT: Beasley, Ellen
; APPLICANT: DiFrancesco, Valentina
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669PCT
; CURRENT APPLICATION NUMBER: US/09/609,816
; CURRENT FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: HUMAN
US-09-609-816-4

Query Match      26.2%; Score 401; DB 2; Length 304;
Best Local Similarity 31.0%; Pred. No. 4.8e-37;
Matches 93; Conservative 62; Mismatches 125; Indels 20; Gaps 5;

Qy 1 MAESEATPTSGEFE-SKYFEFHGVRLPPFCRCGMEEIANFPVRPSDVWIVTPKSGTS 59
Db 1 MAKIEKNAPTMEKKPELFNIMEVDGPTLLISKWEKVCNFOAKPDDLLIATPKSGTT 60
Qy 60 LLQEVVYLVSQADPDEIGLMNIDEQLPVLE--YP---OPGLDIIKELTSPRLIKSHLPY 114
Db 61 WMHEILDMLNDGDVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQLIKTHLPS 120
Qy 115 RFLPSDLHNGDSKVIYMARNPDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGVGS 174
Db 121 HLIPPSIWKENCKIVYVARNPKDCLVSYHFRHMASFMPDPQNLSEFYEKFMVGKVGGS 180
Qy 175 WFEHVQEFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHECHQLV 234
Db 181 WFDHVKGWAAAKMHRILYLFYEDIKKPKHEIHKVLEFLEKTSWGDVINKIVHHTSFDV 240
Qy 235 ---DOCCNAEALP-----VGRGRVGLWKDIFTVSMNEKPDLYVKQMGKCDLTF 280
Db 241 MKDNPMANHTAVPAHIFNHSISKFMKGMGPDGDKNHTVALNENFDKHYEKKMAGSTLNF 300

RESULT 38
US-10-199-334-4
; Sequence 4, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-4

Query Match      26.2%; Score 401; DB 2; Length 304;
Best Local Similarity 31.0%; Pred. No. 4.8e-37;
Matches 93; Conservative 62; Mismatches 125; Indels 20; Gaps 5;

Qy 1 MAESEATPTSGEFE-SKYFEFHGVRLPPFCRCGMEEIANFPVRPSDVWIVTPKSGTS 59
Db 1 MAKIEKNAPTMEKKPELFNIMEVDGPTLLISKWEKVCNFOAKPDDLLIATPKSGTT 60
Qy 60 LLQEVVYLVSQADPDEIGLMNIDEQLPVLE--YP---OPGLDIIKELTSPRLIKSHLPY 114
Db 61 WMHEILDMLNDGDVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQLIKTHLPS 120
Qy 115 RFLPSDLHNGDSKVIYMARNPDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGVGS 174
Db 121 HLIPPSIWKENCKIVYVARNPKDCLVSYHFRHMASFMPDPQNLSEFYEKFMVGKVGGS 180
Qy 175 WFEHVQEFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHECHQLV 234
Db 181 WFDHVKGWAAAKMHRILYLFYEDIKKPKHEIHKVLEFLEKTSWGDVINKIVHHTSFDV 240
Qy 235 ---DOCCNAEALP-----VGRGRVGLWKDIFTVSMNEKPDLYVKQMGKCDLTF 280
Db 241 MKDNPMANHTAVPAHIFNHSISKFMKGMGPDGDKNHTVALNENFDKHYEKKMAGSTLNF 300

RESULT 38
US-10-199-334-4
; Sequence 4, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-4

Query Match      26.2%; Score 401; DB 2; Length 304;
Best Local Similarity 31.0%; Pred. No. 4.8e-37;
Matches 93; Conservative 62; Mismatches 125; Indels 20; Gaps 5;

Qy 1 MAESEATPTSGEFE-SKYFEFHGVRLPPFCRCGMEEIANFPVRPSDVWIVTPKSGTS 59
Db 1 MAKIEKNAPTMEKKPELFNIMEVDGPTLLISKWEKVCNFOAKPDDLLIATPKSGTT 60
Qy 60 LLQEVVYLVSQADPDEIGLMNIDEQLPVLE--YP---OPGLDIIKELTSPRLIKSHLPY 114
Db 61 WMHEILDMLNDGDVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQLIKTHLPS 120
Qy 115 RFLPSDLHNGDSKVIYMARNPDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGVGS 174
Db 121 HLIPPSIWKENCKIVYVARNPKDCLVSYHFRHMASFMPDPQNLSEFYEKFMVGKVGGS 180
Qy 175 WFEHVQEFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHECHQLV 234
Db 181 WFDHVKGWAAAKMHRILYLFYEDIKKPKHEIHKVLEFLEKTSWGDVINKIVHHTSFDV 240
Qy 235 ---DOCCNAEALP-----VGRGRVGLWKDIFTVSMNEKPDLYVKQMGKCDLTF 280
Db 241 MKDNPMANHTAVPAHIFNHSISKFMKGMGPDGDKNHTVALNENFDKHYEKKMAGSTLNF 300

RESULT 39
US-10-199-329-4
; Sequence 4, Application US/10199329
; Patent No. 6953681
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-4

Query Match      26.2%; Score 401; DB 2; Length 304;
Best Local Similarity 31.0%; Pred. No. 4.8e-37;
Matches 93; Conservative 62; Mismatches 125; Indels 20; Gaps 5;

Qy 1 MAESEATPTSGEFE-SKYFEFHGVRLPPFCRCGMEEIANFPVRPSDVWIVTPKSGTS 59
Db 1 MAKIEKNAPTMEKKPELFNIMEVDGPTLLISKWEKVCNFOAKPDDLLIATPKSGTT 60
Qy 60 LLQEVVYLVSQADPDEIGLMNIDEQLPVLE--YP---OPGLDIIKELTSPRLIKSHLPY 114
Db 61 WMHEILDMLNDGDVEKCKRAQTLDRHAFLELKFPHKEKPDLEFVLEMSSPQLIKTHLPS 120
Qy 115 RFLPSDLHNGDSKVIYMARNPDLVSVYQFHRSLRTMSYRGTFQFCRRFMNDKLGVGS 174
Db 121 HLIPPSIWKENCKIVYVARNPKDCLVSYHFRHMASFMPDPQNLSEFYEKFMVGKVGGS 180
Qy 175 WFEHVQEFWEHRMDSNVLFKYEDMHRDLVTMVEQLARFLGVSCDKAQLAETHECHQLV 234
Db 181 WFDHVKGWAAAKMHRILYLFYEDIKKPKHEIHKVLEFLEKTSWGDVINKIVHHTSFDV 240
Qy 235 ---DOCCNAEALP-----VGRGRVGLWKDIFTVSMNEKPDLYVKQMGKCDLTF 280
Db 241 MKDNPMANHTAVPAHIFNHSISKFMKGMGPDGDKNHTVALNENFDKHYEKKMAGSTLNF 300
```

```
Qy 175 WFEHVQEFWEHRMDSNVFLKYEYDMHRLDVTWVEQLARFLGVSCDKAOLEALTEHCHOLV 234
Db 181 WFDHVKGWMAAKDMHRLDLYFYEDIKKPKKEIHKVLEFLEKTSKSGDVINKIVHHTSDV 240
Qy 235 ---DQCCNAEALP-----VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
Db 241 MKDNPMAHNTAVPAHFNHSISKPMKMGPGDKNKHFTVALNENFDKHYEKKMAGSTLNF 300

RESULT 40
US-10-199-330-4
; Sequence 4, Application US/10199330
; Patent No. 6967094
; ORGANISM: Human
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-4
```

```
Query Match 26.2%; Score 401; DB 2; Length 304;
Best Local Similarity 31.0%; Pred. No. 4.8e-37;
Matches 93; Conservative 62; Mismatches 125; Indels 20; Gaps 5;

Qy 1 MAESAETPSPGPEPE-SKYEFHGVRLPPFCRCGMKEEIANFPVRPSDVWIVTPKSGTS 59
Db 1 MAKTEKNAFTWEKPELFNEMVDGVTLLLSKEWKEVCNQAKPDDLILATPKSGTT 60
Qy 60 LLQEVVYLVSQGADPDEIGLWIDQLPVE--YP---QPGLDIIKELTSPRLIKSHLPY 114
Db 61 WHHEITLMDLNDGVEKCKRAQTLDRHAFLELKPPEKPKDLEFVLEMSPLQIKTHLPS 120
Qy 115 RFLPSDLHNGDSKVIYMARNPDLVYVYQFHRSLRTSYRGTFQFCRRFMDKLGYS 174
Db 121 HLIPSIWKNCKIYVYARNKDCLVSYHFRMASFMPDPQNLVEEFYEFKFMGKVGGS 180
Qy 175 WFEHVQEFWEHRMDSNVFLKYEYDMHRLDVTWVEQLARFLGVSCDKAOLEALTEHCHOLV 234
Db 181 WFDHVKGWMAAKDMHRLDLYFYEDIKKPKKEIHKVLEFLEKTSKSGDVINKIVHHTSDV 240
Qy 235 ---DQCCNAEALP-----VGRGRVGLWKDIFTVSMNEKFDLVYKQMGKCDLTF 280
Db 241 MKDNPMAHNTAVPAHFNHSISKPMKMGPGDKNKHFTVALNENFDKHYEKKMAGSTLNF 300
```

```
RESULT 41
US-09-949-016-6526
; Sequence 6526, Application US/09949016
; Patent No. 6812339
; ORGANISM: Human
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6526
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6526
```

```
Query Match 24.9%; Score 381; DB 2; Length 350;
Best Local Similarity 33.3%; Pred. No. 1.1e-34;
Matches 94; Conservative 52; Mismatches 108; Indels 28; Gaps 7;

Qy 12 PGEFESKYEFHGVRLPPFCRG--KMBEIA---NFPVRPSDVWIVTPKSGTSLLOEVV 65
Db 13 PGE-----YFRYKGV---PPFVGLYSLESISLAENTQDVRDDDIIFIITPKSGTTWMIIEI 65
Qy 66 YLVSQGADPDDEIGLWIDQLPVEYQPGLDIIKELTSPRLIKSHLPYRFLPSDLHNGD 125
Db 66 CLILKEGDPWSWIRSVPIWERAPWCE-TIVGAFSLPDQYSPRLMSSHLPIQIFTKAFFSSK 124
Qy 126 SKVIYMARNPDLVYVYQFHRSLRTSYRGTFQFCRRFMDKLGYSWFHVEHVOEFWEH 185
Db 125 AKVIYMGRRPRDVVVSLSYHYSKIAGQLKDPQTPQDFLRDLKGEVQFGSFPDHIKGMILRM 184
Qy 186 RMDSNVLFLKYEDMHRDLVTWVEQLARFLGVSCDKAOLEALTEH---CHQLVDQCCNAEA 242
Db 185 KGDNFLFIITYEELQDQLQGSVERICGFLRPLGKEALGVSVAHSTFSAMKANTMSNTYL 244
Qy 243 LP-----VGRGRVGLWKDIFTVSMNEKFDLVYKQRM 273
Db 245 LPPSLLDHRRGAFRLKGVCGDKNKHFTVAQSEAFDRAYRKQM 286
```

## RESULT 42

```
US-09-949-016-11211
; Sequence 11211, Application US/09949016
; Patent No. 6812339
; ORGANISM: Human
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11211
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11211
```

```
Query Match 24.9%; Score 381; DB 2; Length 377;
Best Local Similarity 33.3%; Pred. No. 1.3e-34;
Matches 94; Conservative 52; Mismatches 108; Indels 28; Gaps 7;

Qy 12 PGEFESKYEFHGVRLPPFCRG--KMBEIA---NFPVRPSDVWIVTPKSGTSLLOEVV 65
Db 40 PGE-----YFRYKGV---PPFVGLYSLESISLAENTQDVRDDDIIFIITPKSGTTWMIIEI 92
Qy 66 YLVSQGADPDDEIGLWIDQLPVEYQPGLDIIKELTSPRLIKSHLPYRFLPSDLHNGD 125
```



```
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: POLYNUCLEOTIDES ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 303
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926-2

Query Match      19.8%; Score 303.5; DB 2; Length 303;
Best Local Similarity 26.7%; Pred. No. 6.6e-26;
Matches 74; Conservative 61; Mismatches 99; Indels 43; Gaps 9;

QY 19 YFEFHGVRLP-PFCRGK-MEEIANFPVRPSDVWIVTPKSGTS-----LLQEVVYLSQGA 72
DB 27 FPTYQGIPYPTMCTSETFQALDTFEARHDDIVLASYPKCGSNWTLHIVSELIVAVSK-- 84
QY 73 DPDEIGLWNIDBQLPVLEYPQ-PGLDI-----IKELTSPRLIKSHLPYRFLPSDLHN 123
DB 85 -----KKYKYPEFPVLECGDSEKQYQRMKGFPSPRILATHLHYDKLPGSIFE 130
QY 124 GDSKVIYMARPKDLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFFHVQBFW 183
DB 131 NKAKILVIFRNPKDTAVSFLHFHNDVDPIDPSYGSWDEFFRQFMKGQVSWGRYFDPAINWN 190
QY 184 EHRMDSNVLFUKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNAEL 243
DB 191 KHLGDGNVKFILDENLAAGIKQIAEFLGFLTGEOITISV---QSTFOAMRAKSQ 247
QY 244 ----PVG-----RGRVGLWKDIFTVSMNEKFDLVYKQ 271
DB 248 DTHGAVGPFPRKGEVGDWKNLFSIQNQEMDEKFE 284

RESULT 46
US-10-364-774-2
; Sequence 2, Application US/10364774
; Patent No. 6929937
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: POLYNUCLEOTIDES ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6

Query Match      19.8%; Score 303.5; DB 2; Length 303;
Best Local Similarity 26.7%; Pred. No. 6.6e-26;
Matches 74; Conservative 61; Mismatches 99; Indels 43; Gaps 9;

QY 19 YFEFHGVRLP-PFCRGK-MEEIANFPVRPSDVWIVTPKSGTS-----LLQEVVYLSQGA 72
DB 27 FPTYQGIPYPTMCTSETFQALDTFEARHDDIVLASYPKCGSNWTLHIVSELIVAVSK-- 84
QY 73 DPDEIGLWNIDBQLPVLEYPQ-PGLDI-----IKELTSPRLIKSHLPYRFLPSDLHN 123
DB 85 -----KKYKYPEFPVLECGDSEKQYQRMKGFPSPRILATHLHYDKLPGSIFE 130
QY 124 GDSKVIYMARPKDLVSYQFHRSLRTMSYRGTFQFCRRFMNDKLGYSWFFHVQBFW 183
DB 131 NKAKILVIFRNPKDTAVSFLHFHNDVDPIDPSYGSWDEFFRQFMKGQVSWGRYFDPAINWN 190
QY 184 EHRMDSNVLFUKYEDMHRDLVTWVEQLARFLGVSCDKAQLAEALTEHCHQLVDQCCNAEL 243
DB 191 KHLGDGNVKFILDENLAAGIKQIAEFLGFLTGEOITISV---QSTFOAMRAKSQ 247
QY 244 ----PVG-----RGRVGLWKDIFTVSMNEKFDLVYKQ 271
DB 248 DTHGAVGPFPRKGEVGDWKNLFSIQNQEMDEKFE 284

RESULT 47
US-09-795-926-6
; Sequence 6, Application US/09795926
; Patent No. 6555669
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: POLYNUCLEOTIDES ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
```



```
; LENGTH: 265
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926-6

Query Match      19.1%; Score 292.5; DB 2; Length 265;
Best Local Similarity 27.1%; Pred. No. 9.7e-25;
Matches 70; Conservative 57; Mismatches 90; Indels 41; Gaps 7;

QY 36 EEIANFPVPSDWMIVTPKSGTS-----LLQEVVYLVSQADPDEIGLNMIDQLPVLRY 91
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 8 QALDTFEARHDDIVLASYPKCGSNWILHIVSELIIYAVSK-----KKYKY 51
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 92 PQ-PGLDI-----IKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDLVVSF 142
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 52 BEFPVLECGDSEKQYRMKGFPSPRILATHLHYDKLPGSIFENKAKILVIFRNPKDTAVSF 111
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 143 YQFHSRLRTMSVGTQFCRCRFMNDKLGYSWFHVEHWRMDSNVLFUKYEDMHRD 202
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 112 LHFHNDVPDIPSYGSDWDEFFRQFMKGQVSWGRYFDFAINWNKHLGDGDNVKFILDLEN 171
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 203 LVTMVEQLARFLGVSCDKAQLALTEHCHQLVDQCCNAEAL-----PVG-----RGRVGLW 253
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 172 LAAGIKQIAEFLGFFLTGEGIQTISV---QSTFQAMRAKSQDTHGAVGFLFRKGEVGDW 228
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 254 KDIFTVSMNEKEFDLVYKQ 271
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 229 KNLFSEIQNQEWDKEFKE 246
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 48
US-10-364-774-6
; Sequence 6, Application US/10364774
; Patent No. 6929937
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: POLYNUCLEOTIDES ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/10/364,774
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 265
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774-6

Query Match      19.1%; Score 292.5; DB 2; Length 265;
Best Local Similarity 27.1%; Pred. No. 9.7e-25;
Matches 70; Conservative 57; Mismatches 90; Indels 41; Gaps 7;

QY 36 EEIANFPVPSDWMIVTPKSGTS-----LLQEVVYLVSQADPDEIGLNMIDQLPVLRY 91
```

```
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 8 QALDTFEARHDDIVLASYPKCGSNWILHIVSELIIYAVSK-----KKYKY 51
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 92 PQ-PGLDI-----IKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARPNKDLVVSF 142
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 52 BEFPVLECGDSEKQYRMKGFPSPRILATHLHYDKLPGSIFENKAKILVIFRNPKDTAVSF 111
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 143 YQFHSRLRTMSVGTQFCRCRFMNDKLGYSWFHVEHWRMDSNVLFUKYEDMHRD 202
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 112 LHFHNDVPDIPSYGSDWDEFFRQFMKGQVSWGRYFDFAINWNKHLGDGDNVKFILDLEN 171
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 203 LVTMVEQLARFLGVSCDKAQLALTEHCHQLVDQCCNAEAL-----PVG-----RGRVGLW 253
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 172 LAAGIKQIAEFLGFFLTGEGIQTISV---QSTFQAMRAKSQDTHGAVGFLFRKGEVGDW 228
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 254 KDIFTVSMNEKEFDLVYKQ 271
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 229 KNLFSEIQNQEWDKEFKE 246
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 49
US-09-854-122-20
; Sequence 20, Application US/09854122
; Patent No. 6841718
; GENERAL INFORMATION:
; APPLICANT: ALBERTE, RANDALL S.
; APPLICANT: SMITH, ROBERT
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
; FILE REFERENCE: PHA-007.01
; CURRENT APPLICATION NUMBER: US/09/854,122
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/202,529
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Flaveria bidentis
US-09-854-122-20
```

```
Query Match      18.0%; Score 275.5; DB 2; Length 320;
Best Local Similarity 27.5%; Pred. No. 1.1e-22;
Matches 80; Conservative 63; Mismatches 111; Indels 37; Gaps 11;

QY 18 KYFEFHGVLPPFCRGKMEETANFPVPSDWMIVTPKSGTSLLOEVVYLVSQADPDSI 77
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 34 KYQDFWG--LQNNIEGAILAQOSFKARDDVFLCSYPKSGTTWLKALAYAVITREKDFEF 91
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 78 G---LMNIDEQ-LPVLEYPPQGLDIIKELTSPR-----LIKSHLPYRFLPSDLHNGDSKV 128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 92 TSPLLTNIPHCIPYIE-----KDLKKIVENQNNSCFTPMATHMPYHVLPKSILALNCKM 146
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 129 IYMARPNKDLVVSYYQFHSRLRTMSYR-GTQEFQRCRRMNDKLGYSWFHVEHWRM 187
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 147 VYIYRNKDVIVSYFHYFGREITKLPLEDAFEEAFDEFYHGISQFGPYWDHLLGYWKASL 206
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 188 D--SNVLFUKYEDMHRDLVTMVEQLARFLGVSC-----DKAOLEALTEHC-----HQVL 234
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 207 ERPEVILFKYEDVKKDPTSNVKRLAEFIGYPTTPEEKEGVIESIIKLCSPENLSNLEV 266
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 235 DOCCNAEA-LPV-----GRRVGLWKKDIFTVSMNEKEFDLVYKQRMGKCDL 278
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 267 NKSGNSKGFLPIENLYFRKAKGDWKNYFTDEMTEDIKLDKLSATGL 317
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 50
US-09-854-122-18
; Sequence 18, Application US/09854122
; Patent No. 6841718
; GENERAL INFORMATION:
; APPLICANT: ALBERTE, RANDALL S.
; APPLICANT: SMITH, ROBERT
```

```
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
; FILE REFERENCE: PHA-007.01
; CURRENT APPLICATION NUMBER: US/09/854,122
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/202,529
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Brassica napus
US-09-854-122-18

Query Match      17.4%; Score 265.5; DB 2; Length 324;
Best Local Similarity 26.8%; Pred. No. 1.6e-21;
Matches 80; Conservative 61; Mismatches 128; Indels 29; Gaps 11;

QY  9 PSTGFEFSKYFEFHG-VRLPPFCRGKMBEIANFFVRPSDVMIVTPKSGTSLLOEVVY-
DB 28 PSEKGLVCMQVQFQGRWHTQALLQILTCQHFKAQSDIILVTNPKSGTTWLKALVFA 87
QY 67 LVSOAGDPDEIGLNMIDQLPVLVEYQPGDLIIK-----ELTSPRLIKSHLPYRFLPS 119
DB 88 LINRHKFP-VYSVIIISCVQSALLVPFLGRSLRSPDFDFSQSSPRLMNTHTSHLSLPE 146
QY 120 DLHNGDSKVIYARNPKDLVSYVQFHRSL----RTMSYRGTFQFRCRRPMNDKLGYSWFW 176
DB 147 SVKSSCKIVYCCRNPKDMFVSLWHFGKLAPEETADY--PIEKAVEAFCEGKFGPW 204
QY 177 EHVQFWEHRMD--SNVLFLKYEDMHRDLVTMVEQLARFLGVS-CDKAQLEALTEHC--H 231
DB 205 DHVLSYTWASLNPKNKVLFSVSEEPKKTGETIKRIAEFLGGLGVLGEVEVRAIVKLCSE 264
QY 232 QLVDOCCNAEA-LPVG-----RGRVGLMKDIFTVSMNEKFDLVYKQKMGKCDLTF 280
DB 265 SLSSLEVNREGKLPSCMETRAFFRKGVEGWEDTLTESLAEVIDRTIEBKFGGSLKF 322

RESULT 51
US-09-854-122-19
; Sequence 19, Application US/09854122
; Patent No. 6841718
; GENERAL INFORMATION:
; APPLICANT: ALBERTE, RANDALL S.
; APPLICANT: SMITH, ROBERT
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
; FILE REFERENCE: PHA-007.01
; CURRENT APPLICATION NUMBER: US/09/854,122
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/202,529
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-854-122-19

Query Match      17.0%; Score 259.5; DB 2; Length 302;
Best Local Similarity 26.9%; Pred. No. 7e-21;
Matches 77; Conservative 52; Mismatches 108; Indels 49; Gaps 12;

QY  9 PSTGFEFSKYFEFHG-VRLPPFCRGKMBEIANFFVRPSDVMIVTPKSGTSLLOEVVYL 67
DB 28 PREKGLVSEIVEFQGLWHTQAILQGLICQKRFKAQSDIILVTNPKSGTTWLKALVFA 87
QY 68 V-----SQADPDEIGLNMIDQLPVL---YPCGLDIIKELTSPRLIKSHLPYR 115
DB 88 LINRHKFPVSSSGNHP--LLVTNPHLLVPFLGVYVESPDFD-FSSLPSRLMNTHTSHL 144
QY 116 FLPSDLHNGDSKVIYARNPKDLVSYVQFHRSL---RTMSYRGTFQFRCRRFMNDKLG 172
```

```
DB 145 SLPSVSKSSCKIYVCCRNPKDMFVSLWHFGKLAPEETADY--PIEKAVEAFCEGKFG 202
QY 173 GSWFEHVQEFW--BHRMDSNVFLKYEDMHRDLVTMVEQLARFLGVSQCDKAQLEALTEHC 230
DB 203 GPFWDHILEYWIYASRENPNKVLFTYBELKQTEVEMKRIAETL--ECGFIESBEVRE-- 258
QY 231 HOLVDQCCNAEA-----LPVG-----RGRVGLMKDIF 257
DB 259 ---IVKLCSPESLSNLEVNKEGKLPNGIETKTFPRKGEIGGWRDSF 301

RESULT 52
US-09-795-926-16
; Sequence 16, Application US/09795926
; Patent No. 6555869
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kiese, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 214
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926-16

Query Match      16.6%; Score 254; DB 2; Length 214;
Best Local Similarity 30.2%; Pred. No. 1.8e-20;
Matches 55; Conservative 42; Mismatches 73; Indels 12; Gaps 3;

QY  99 IKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYARNPKDLVSYVQFHRSLRTWSYGTGF 158
DB 17 MKGFPSPRILATHLHYDKLPGSIFENKAKILVIFRNPKDTAVSFHLHFHNDVPDIPSYGSM 76
QY 159 QEFCCRFRMNDKLGYSWFEHVQEFWEHRMDSNVFLKYEDMHRDLVTMVEQLARFLGVSC 218
DB 77 DEFRQFMKGQVSWGRYDFDAINWNKHLDDGNVXFILYEDLKENLAAGIKQIAEFLGPFLL 136
QY 219 DKAQLEALTEHCQLVDQCCNAEAAL----PVG-----RGRVGLMKDIFTVSMNEKFDLVY 269
DB 137 TGEQIQTISV---QSTFOAMRAKSQDTHGAVGPFPRKGEVGDWKNLFSIQNDEMDEKF 193
QY 270 KQ 271
DB 194 KE 195

RESULT 53
US-10-364-774-16
; Sequence 16, Application US/10364774
; Patent No. 6929937
```





```
Best Local Similarity 27.5%; Pred. No. 8.3e-12;
Matches 42; Conservative 31; Mismatches 50; Indels 30; Gaps 5;

Qy 36 BEIANFPVRPSDVWIVTPKSGTS-----LLOEVVVLVSQGADPDPEIGLWNIDEQLPVLEY 91
Db 8 QALDTFEARHDDIVLASYPKCGSNWLHIVSELIVAVSK-----KKYKY 51

Qy 92 PQ-PGLDI-----IKELTSPRLKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVVSY 142
Db 52 PEFPVLECGDSEKYQRMKGFPSPRILATHLHYDKLPGSIFENKAKILVIFRNPKDTAVSF 111

Qy 143 YQFHRSLRTMSYRGTFQFCFRFNMNDKLGYSW 175
Db 112 LHFHNDVPDIPSYGSWDEFFRQFMKGQS-GCW 143
```

```
RESULT 59
US-10-364-774-8
; Sequence 8, Application US/10364774
; Patent No. 6929937
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/10/364,774
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 148
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774-8
```

```
Query Match 11.5%; Score 176; DB 2; Length 148;
Best Local Similarity 27.5%; Pred. No. 8.3e-12;
Matches 42; Conservative 31; Mismatches 50; Indels 30; Gaps 5;

Qy 36 BEIANFPVRPSDVWIVTPKSGTS-----LLOEVVVLVSQGADPDPEIGLWNIDEQLPVLEY 91
Db 8 QALDTFEARHDDIVLASYPKCGSNWLHIVSELIVAVSK-----KKYKY 51

Qy 92 PQ-PGLDI-----IKELTSPRLKSHLPYRFLPSDLHNGDSKVIYMARNPKDLVVSY 142
Db 52 PEFPVLECGDSEKYQRMKGFPSPRILATHLHYDKLPGSIFENKAKILVIFRNPKDTAVSF 111

Qy 143 YQFHRSLRTMSYRGTFQFCFRFNMNDKLGYSW 175
Db 112 LHFHNDVPDIPSYGSWDEFFRQFMKGQS-GCW 143
```

```
RESULT 60
US-10-094-749-2841
```

```
; Sequence 2841, Application US/10094749
; Patent No. 6979557
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2841
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2841
```

```
Query Match 9.8%; Score 150.5; DB 2; Length 106;
Best Local Similarity 37.2%; Pred. No. 4.1e-09;
Matches 32; Conservative 18; Mismatches 33; Indels 3; Gaps 1;

Qy 36 BEIANFPVRPSDVWIVTPKSGTSLLQEVVVLVSQGADPDPEIGLWNIDEQLPVLEY---P 92
Db 6 DKIWNFQAKPDLLIASYPKAGITWTQEIQLDQNGDIEKSRASIQLOHPFLEWTRMT 65

Qy 93 QPGLDIIEKLTSPRLIKSHLPYRFLP 118
Db 66 HCGIDQANTWPSRPTLTKHLPVQLLP 91
```

```
RESULT 61
US-09-795-926-18
; Sequence 18, Application US/09795926
; Patent No. 6555669
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
```

; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 60/186,558  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: US 60/191,849  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 18  
; LENGTH: 97  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-795-926-18

Query Match 9.0%; Score 137.5; DB 2; Length 97;  
Best Local Similarity 35.1%; Pred. No. 1.1e-07;  
Matches 27; Conservative 16; Mismatches 33; Indels 1; Gaps 1;  
Qy 99 IKELTSPLIKSHLYRFLPSDLHNGDSKVIYMARPKDLVVSYQFHRSLRTMSYRGTG 158  
Db 17 MKGFSPRILATHLHYDKLPGSIFENKAKILVIFRNPKDTAVSFLHFHNDVPDIPSYGSW 76  
Qy 159 QFCRRFNMNKLGYGSW 175  
Db 77 DEFRRQFMKGQES-GCW 92

RESULT 62  
US-10-364-774-18  
; Sequence 18, Application US/10364774  
; Patent No. 6929937  
; GENERAL INFORMATION:  
; APPLICANT: Donoho, Gregory  
; APPLICANT: Hilbun, Erin  
; APPLICANT: Turner, C. Alexander Jr.  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Abuin, Alejandro  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Sands, Arthur T.  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Wilganowski, Nathaniel L.  
; APPLICANT: Hu, Yi  
; APPLICANT: Kieke, James Alvin  
; APPLICANT: Potter, David George  
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND  
; FILE REFERENCE: POLYNUCLEOTIDES ENCODING THE SAME  
; CURRENT APPLICATION NUMBER: US/10/364,774  
; CURRENT FILING DATE: 2003-02-11  
; PRIOR APPLICATION NUMBER: US/09/795,926  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 60/185,920  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 60/186,558  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: US 60/191,849  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 18  
; LENGTH: 97  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-364-774-18

Query Match 9.0%; Score 137.5; DB 2; Length 97;  
Best Local Similarity 35.1%; Pred. No. 1.1e-07;  
Matches 27; Conservative 16; Mismatches 33; Indels 1; Gaps 1;  
Qy 99 IKELTSPLIKSHLYRFLPSDLHNGDSKVIYMARPKDLVVSYQFHRSLRTMSYRGTG 158  
Db 17 MKGFSPRILATHLHYDKLPGSIFENKAKILVIFRNPKDTAVSFLHFHNDVPDIPSYGSW 76  
Qy 159 QFCRRFNMNKLGYGSW 175

Db 77 DEFRRQFMKGQES-GCW 92  
RESULT 63  
US-09-513-999C-8076  
; Sequence 8076, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59. US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 8076  
; LENGTH: 104  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 9  
; OTHER INFORMATION: Xaa=Asp or Glu  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 12  
; OTHER INFORMATION: Xaa=Glu or Gly  
US-09-513-999C-8076

Query Match 7.5%; Score 115; DB 2; Length 104;  
Best Local Similarity 31.2%; Pred. No. 4.5e-05;  
Matches 24; Conservative 18; Mismatches 35; Indels 0; Gaps 0;  
Qy 14 EFESKYFEFHGVRLPPFCRGMBEIANFPVRPSDVWITYPKSGTSLQLQEVVYLVSGAD 73  
Db 6 DYXKFXEVHGILMYKDFVYWDNVFAQARPDLLVIATYPKSGTTWVSEIVVMYIKEGD 65  
Qy 74 PDEIGLNMNIDQLPVLE 90  
Db 66 VEKCKEDVIFNRIPFLE 82

RESULT 64  
US-09-167-681-49  
; Sequence 49, Application US/09167681A  
; Patent No. 6265561  
; GENERAL INFORMATION:  
; APPLICANT: Weinshilboum, M.D., Richard M.  
; APPLICANT: Raftogianis, Rebecca B.  
; APPLICANT: Wood, Thomas C.  
; APPLICANT: Otterness, Diane M.  
; TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS  
; FILE REFERENCE: 07039/118001  
; CURRENT APPLICATION NUMBER: US/09/167,681A  
; CURRENT FILING DATE: 1998-10-07  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 49  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-167-681-49

Query Match 7.3%; Score 111; DB 2; Length 42;  
Best Local Similarity 51.3%; Pred. No. 3.3e-05;  
Matches 20; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

```
Qy 128 VYIARNPKDLVSVYQFHSRLRMTSYRGTFQFQFRRFM 166
Db 1 VYIARNPKDLVSVYQFHSRLRMTSYRGTFQFQFRRFM 166
; PRIORITY: 128
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-768-158-2.ra1

Query Match 7.0%; Score 106.5; DB 2; Length 325;
Best Local Similarity 23.7%; Pred. No. 0.0024;
Matches 65; Conservative 37; Mismatches 105; Indels 67; Gaps 14;

Qy 43 VRPSDWIVTPKSGTSLQEVV-YLVSGAD-PDEIGLMN--IDEQLPVLEYPOPGLDI 98
Db 27 LRGDIIISAPSKSLGTQRLVSLVFDGDLPGPLSTVSPWLDQTRPIEEVVATLDA 86
Qy 99 IKELTSPLIKSHLPYRFLPSDLHNGDSKVY--MARNPKDLVSV-YVQ----- 144
Db 87 QQH---RRFIKTHPLDGLV-----DDRSYICVGRDPRDAAVSMLYQSANNMNRMI 138
Qy 145 -----FHSRL-----RTMSYRGTFQFQFRRFM--DKLGYGSMFHEVQ----- 180
Db 139 LHEAVVPFHERIAPPFAELGHARSPTTEF-RDWMEGNQPPPGIGFTHLKGIGTLANTLH 197
Qy 181 ---EFWEHRMDSNVFLKYEDMHRDLVTMVQLARFLGVSCDK-----AQLEALTE 228
Db 198 QLGTWVVRHLPNVALFYADYQADLAGELLRPARVLGIAATRDARDLAQYATLDAMS 257
Qy 229 HCHQL-----VDQCCNAEALPVGRGRVGLWKDIFT 258
Db 258 RASEIAPNTTGTGHWSDERFRFRGGSGDWQOFFT 291

RESULT 67
US-10-768-158-2.ra1
; Sequence 25, Application US/10891383A
; Patent No. 6974580
; GENERAL INFORMATION:
; APPLICANT: Bertozzi, Carolyn
; APPLICANT: Williams, Spencer J.
; APPLICANT: Mougous, Joseph
; TITLE OF INVENTION: Mycobacterial Sulfation Pathway Proteins
; FILE REFERENCE: and Methods of Use Thereof
; CURRENT APPLICATION NUMBER: US/10/891,383A
; CURRENT FILING DATE: 2004-07-13
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/285,394
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: 10/126,279
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-768-158-2.ra1

Query Match 7.0%; Score 106.5; DB 2; Length 325;
Best Local Similarity 23.7%; Pred. No. 0.0024;
Matches 65; Conservative 37; Mismatches 105; Indels 67; Gaps 14;

Qy 43 VRPSDWIVTPKSGTSLQEVV-YLVSGAD-PDEIGLMN--IDEQLPVLEYPOPGLDI 98
Db 28 LRGDIIISAPSKSLGTQRLVSLVFDGDLPGPLSTVSPWLDQTRPIEEVVATLDA 87
Qy 99 IKELTSPLIKSHLPYRFLPSDLHNGDSKVY--MARNPKDLVSV-YVQ----- 144
Db 88 QQH---RRFIKTHPLDGLV-----DDRSYICVGRDPRDAAVSMLYQSANNMNRMI 139
Qy 145 -----FHSRL-----RTMSYRGTFQFQFRRFM--DKLGYGSMFHEVQ----- 180
Db 139 LHEAVVPFHERIAPPFAELGHARSPTTEF-RDWMEGNQPPPGIGFTHLKGIGTLANTLH 197
Qy 181 ---EFWEHRMDSNVFLKYEDMHRDLVTMVQLARFLGVSCDK-----AQLEALTE 228
Db 198 QLGTWVVRHLPNVALFYADYQADLAGELLRPARVLGIAATRDARDLAQYATLDAMS 257
Qy 229 HCHQL-----VDQCCNAEALPVGRGRVGLWKDIFT 258
Db 258 RASEIAPNTTGTGHWSDERFRFRGGSGDWQOFFT 291

RESULT 66
US-10-768-158-2.ra1
; Sequence 25, Application US/10286606
; Patent No. 6863895
; GENERAL INFORMATION:
; APPLICANT: Bertozzi, Carolyn
; APPLICANT: Williams, Spencer J.
; APPLICANT: Mougous, Joseph
; TITLE OF INVENTION: Mycobacterial Sulfation Pathway Proteins
; FILE REFERENCE: and Methods of Use Thereof
; CURRENT APPLICATION NUMBER: US/10/286,606
; CURRENT FILING DATE: 2002-10-31
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/285,394
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: 10/126,279
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-768-158-2.ra1
```

Qy 145 -----FHRSL-----RTMSYRGTFQFCRRFMN--DKLGYGSWFHFVQ----- 180  
Db 140 LHEAVVPHERIAPPFAELGHARPTTEF-RDWMEGPNQPPPGIGFTHLKGIGTLANTLH 198  
Qy 181 ---EFWEHRMDSNVLFKYEDHMRDLVTMVQLARFLGVSCDK-----AQLEALTE 228  
Db 199 QLGTWVRRLPNVALPHYADQADLACELLRPARVLGIAATRRDRDLAQVATLDAMRS 258  
Qy 229 HCHQL-----VQCCNAEALPVGRGRVGLWKDIFT 258  
Db 259 RASEIAPNTDGIWHSDEFRFRGGSGDMQOQFFT 292

## RESULT 68

US-09-795-926-10

; Sequence 10, Application US/09795926

; Patent No. 6555669

; GENERAL INFORMATION:

; APPLICANT: Donoho, Gregory

; APPLICANT: Hilbun, Erin

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Friedrich, Glenn

; APPLICANT: Abuin, Alejandro

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Sands, Arthur T.

; APPLICANT: Walke, D. Wade

; APPLICANT: Wilganowski, Nathaniel L.

; APPLICANT: Hu, Yi

; APPLICANT: Kieke, James Alvin

; APPLICANT: Potter, David George

; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND

; FILE REFERENCE: LEX-0144-USA

; CURRENT FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: US/09/795,926

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: US 60/185,920

; PRIOR FILING DATE: 2000-02-29

; PRIOR APPLICATION NUMBER: US 60/186,558

; PRIOR FILING DATE: 2000-03-02

; PRIOR APPLICATION NUMBER: US 60/191,849

; PRIOR FILING DATE: 2000-03-24

; NUMBER OF SEQ ID NOS: 47

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10

; LENGTH: 148

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-795-926-10

Query Match 6.8%; Score 104.5; DB 2; Length 148;  
Best Local Similarity 22.6%; Pred. No. 0.0012;  
Matches 30; Conservative 28; Mismatches 44; Indels 31; Gaps 6;

Qy 19 YFEFHGVRLP-PFCRGK-MEEIANPPVRPSDVWIVTPYKSGTS-----LLOEVVYLVSOQA 72  
Db 27 PFTYQGIPIPTMCTSETFQALDTFEARHDDIVLASYPKCGSNWLHIVSELIVAVSK-- 84  
Qy 73 DPDEIGLWNIDEQLPVLEYPQ-PGLDI-----IKELTSPLRIKSHLPYRFLPSDLHN 123  
Db 85 -----KKYKYPEFPVLECGDSBKYQRMKGFPSPRILATHLHYDKLPGSIFE 130

Qy 124 GDSKVIYMARNP 136

Db 131 NKAKRQHLTMLPR 143

## RESULT 69

US-10-364-774-10

; Sequence 10, Application US/10364774

; Patent No. 6929937

; GENERAL INFORMATION:

; APPLICANT: Donoho, Gregory

; APPLICANT: Hilbun, Erin

; APPLICANT: Turner, C. Alexander Jr.  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Abuin, Alejandro  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Sands, Arthur T.  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Wilganowski, Nathaniel L.  
; APPLICANT: Hu, Yi  
; APPLICANT: Kieke, James Alvin  
; APPLICANT: Potter, David George  
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND  
; FILE REFERENCE: LEX-0144-USA  
; CURRENT APPLICATION NUMBER: US/10/364,774  
; CURRENT FILING DATE: 2003-02-11  
; PRIOR APPLICATION NUMBER: US/09/795,926  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 60/185,920  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 60/186,558  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: US 60/191,849  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 148  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-364-774-10

Query Match 6.8%; Score 104.5; DB 2; Length 148;  
Best Local Similarity 22.6%; Pred. No. 0.0012;  
Matches 30; Conservative 28; Mismatches 44; Indels 31; Gaps 6;

Qy 19 YFEFHGVRLP-PFCRGK-MEEIANPPVRPSDVWIVTPYKSGTS-----LLOEVVYLVSOQA 72  
Db 27 PFTYQGIPIPTMCTSETFQALDTFEARHDDIVLASYPKCGSNWLHIVSELIVAVSK-- 84  
Qy 73 DPDEIGLWNIDEQLPVLEYPQ-PGLDI-----IKELTSPLRIKSHLPYRFLPSDLHN 123  
Db 85 -----KKYKYPEFPVLECGDSBKYQRMKGFPSPRILATHLHYDKLPGSIFE 130  
Qy 124 GDSKVIYMARNP 136  
Db 131 NKAKRQHLTMLPR 143

## RESULT 70

US-09-125-635-12  
; Sequence 12, Application US/09125635  
; Patent No. 6562589

; GENERAL INFORMATION:

; APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE

; TITLE OF INVENTION: AIB1, A novel steroid receptor co-activator

; FILE REFERENCE: 49944

; CURRENT APPLICATION NUMBER: US/09/125,635

; CURRENT FILING DATE: 1998-08-21

; PRIOR APPLICATION NUMBER: 60/049,728

; PRIOR FILING DATE: 1997-06-17

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 12

; LENGTH: 1402

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-125-635-12

Query Match 6.8%; Score 104.5; DB 2; Length 1402;  
Best Local Similarity 21.0%; Pred. No. 0.038;  
Matches 49; Conservative 45; Mismatches 104; Indels 35; Gaps 10;

Qy 14 EFESKYFEFHGVRLPFCRGKMEIANPPVRPSDVWIVTPYKSGTSLLOEVVYLVSOQAD 73







```

; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: NOVEL NUCLEOTIDES ENCODING THE SAME
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 110
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926--4

Query Match          6.1%; Score 93.5; DB 2; Length 110;
Best Local Similarity 22.8%; Pred. No. 0.014;
Matches 26; Conservative 24; Mismatches 35; Indels 29; Gaps 4;

QY 36 EEIANPVRPSDVWIVTPKSGTS---LQEEVVYLVSQGADPDEIGLWNIDEQLPVLEY 91
DB 8 QALDTFARHDDIVLASYPKGSNWILHVSELIYAVSK-----KKYKY 51
QY 92 PQ-PGLDI-----IKELTSPRLKSHLPYRFLPSDLHGNSKVIYMARNPK 136
DB 52 PEPFVLECGDSEKYQRMKGFPSPRILATHLHYDKLPGSIFENKAKRQHLTMLPR 105

RESULT 82
US-10-364-774-4
; Sequence 4, Application US/10364774
; Patent No. 6929937
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/10/364,774
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 110
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774-4

```



Db 729 FYQVISASOTPLALRSLQNRCLVPGYGYSTHLQRLWLTYPSPGQLLIVDQGLRTNPAASM 788

QY 208 EQLAREFLGVS 217  
| : : : : :  
Db 789 ESIQKFLGIT 798

RESULT 86  
US-09-949-016-7255  
; Sequence 7255, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7255  
; LENGTH: 892  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-7255

Query Match 6.0%; Score 92.5; DB 2; Length 892;  
Best Local Similarity 20.8%; Pred. No. 0.45;  
Matches 52; Conservative 37; Mismatches 104; Indels 57; Gaps 11;

QY 12 PGFEFSKYFEHGVRLPPFCRGKMBEIANFPVRPSDVW-----IVTPKSGT 58  
| : : : : :  
Db 571 PVPQAQYFELFPQERSPLWQNPCCD-----KRHKDINSKCTCDRLPKFLVGPQKTGT 625

QY 59 SLLEQVYVL-----VSQADP---DEIGLNN-----IDEQLPVLEYP-QPGLDIIKELT 103  
| : : : : :  
Db 626 TAIHFSLUHPAVTSFSPSTFEIQTFFNSYHKGIDWYMDFFVPVPSNASTDFLFEKS 685

QY 104 SPRLIKSHLPYR---FLPSDLHNGDSKVIYMARNPDKLVVSYVYQFHRSLR---TMSYRGT 157  
| : : : : :  
Db 686 ATYFDEVVPRGAALLPR-----AKIITVLTNPADRAYSVYQORAHGDPVALNY--T 737

QY 158 PQEF-----CRRFNNDKLGYSWFEHVOEFWEHRMDSNVLFKYEDMHRDLVTMV 207  
| : : : : :  
Db 738 FYQVISASOTPLALRSLQNRCLVPGYGYSTHLQRLWLTYPSPGQLLIVDQGLRTNPAASM 797

QY 208 EQLAREFLGVS 217  
| : : : : :  
Db 798 ESIQKFLGIT 807

RESULT 87  
US-09-270-767-41537  
; Sequence 41537, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 41537  
; LENGTH: 121  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster

US-09-270-767-41537

Query Match 5.8%; Score 88; DB 2; Length 121;  
Best Local Similarity 26.0%; Pred. No. 0.07;  
Matches 38; Conservative 20; Mismatches 32; Indels 56; Gaps 9;

QY 78 GLMNIDEQLPVLEYQPGGLDIIKELTSPRLIKSHLPYRFLPSDLHNGDSKVIYMARNPDK 137  
| : : : : :  
Db 14 GLGRADEHVALSQYPRPGM-----LPTKHPNSDKPKL---RNK- 49

QY 138 LVVSYVQFHRSLRTMSYRGTF-----QEFCCRFRMNDKLGYSWFE-HVQEFWEHRMDSN 190  
| : : : : :  
Db 50 -IMGY-----SURTDIYRYTMVRFRHAQNFSR-----DWHDVYGEELYDHRDLDSG 93

QY 191 -----VLFLEKYED---MHRDLVTMV 207  
| : : : : :  
Db 94 BELNLVPLPQFDDVRQRLRRRLMENV 119

RESULT 88  
US-09-949-016-8271  
; Sequence 8271, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8271  
; LENGTH: 610  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-8271

Query Match 5.8%; Score 88; DB 2; Length 610;  
Best Local Similarity 21.7%; Pred. No. 0.82;  
Matches 61; Conservative 38; Mismatches 92; Indels 90; Gaps 15;

QY 21 BFHGVRLPPFCRGKMBEIANFPVRPSDVWIVTPYKSGTSLLOEVVYVVSQADPDEIGLM 80  
| : : : : :  
Db 170 BFH-----SKLQEATNFPRLP---FVIPFLKANLPQLQRELLHCLAKQNPAYL 217

QY 81 NIDEOL-----PV-----LEYQPGLDIIKELTSPRLIK-----SHLP 113  
| : : : : :  
Db 218 AQHEQLLLDASTSPVDSSELLDVNENG---KRTPTDKENGDFDREPLHSEHPSKRP 273

QY 114 YRFLPSDLHNGDSKVIYMARNP-----PK-----DLVVSY-----YQFHRSL----- 149  
| : : : : :  
Db 274 CTISPGQRYSPNNGLSYQPNGLPHTPPPPQHYRLDDMAIAHHYRDSYRHPSHRDLRDN 333

QY 150 RTMSYRGTGFQFCRRFMNDKLGYSWFEHVOEFWEHRMDSNVLFKYEDMHRDLVTMV 209  
| : : : : :  
Db 334 RPMGLHGTRQE---EMIDHRLTDREWAE---WKH-----LDHLLNCIMDWVEK 376

QY 210 LARFLGV--SCDKAQLEALTECHQLVDQCCNAEALPVGRG 248  
| : : : : :  
Db 377 TRRSLTVLRCQCEADREELNYWIRYS-----AEDLKKGGG 413

RESULT 89  
US-09-949-016-8270  
; Sequence 8270, Application US/09949016  
; Patent No. 6812339

```

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8270
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8270

Query Match      5.8%; Score 88; DB 2; Length 638;
Best Local Similarity 21.7%; Pred. No. 0.88;
Matches 61; Conservative 38; Mismatches 92; Indels 90; Gaps 15;

QY 21 EFHGVRLPPFCRGKMEETANFVRPSDVWIVTPKSGTSLLOEVVLYSQGADPDEIGLM 80
DB 198 EFH-----SKLQEATNFPRLP---FVIFPKANLPLLORELLHCARLAKONPAOYL 245
QY 81 NIDEQL-----PV-----LEYQPGLDIIKELTSPRIK-----SHLP 113
DB 246 AQHEQLLDASTTSPVDSSELLDVNENG-----KRTPTDKENGDFREPLHSEHPKRP 301
QY 114 YRFLPSDLHNGDSKVIYARN-----PK-----DLVVS-----YQFHRSL----- 149
DB 302 CTISPGQKRYSPNGLSYQPNGLPHPTPPPPQHYRLDDMAIAHHYRDSYRHPSHRDLDRN 361
QY 150 RTMSYRGTFQFCRRFPMNDKLGYSWFHVEHQRDMSNVFLFKYEDMHRDLVTWVEQ 209
DB 362 RPMGLHGTRQE---EMIDHRLTDREWABE---WKH-----LDHLLNCIMDMVEK 404
QY 210 LARFLGV--SCDKAQLEALTECHQLVDQCCNAEALPVGRG 248
DB 405 TRRSITVLRRCQEADEELNYWIRYS-----AEDLKGGG 441

RESULT 90
US-09-949-016-7288
; Sequence 7288, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7288
; LENGTH: 702
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7288

Query Match      5.8%; Score 88; DB 2; Length 702;
Best Local Similarity 21.7%; Pred. No. 1;
Matches 61; Conservative 38; Mismatches 92; Indels 90; Gaps 15;

QY 21 EFHGVRLPPFCRGKMEETANFVRPSDVWIVTPKSGTSLLOEVVLYSQGADPDEIGLM 80
DB 198 EFH-----SKLQEATNFPRLP---FVIFPKANLPLLORELLHCARLAKONPAOYL 245
QY 81 NIDEQL-----PV-----LEYQPGLDIIKELTSPRIK-----SHLP 113
DB 246 AQHEQLLDASTTSPVDSSELLDVNENG-----KRTPTDKENGDFREPLHSEHPKRP 301
QY 114 YRFLPSDLHNGDSKVIYARN-----PK-----DLVVS-----YQFHRSL----- 149
DB 302 CTISPGQKRYSPNGLSYQPNGLPHPTPPPPQHYRLDDMAIAHHYRDSYRHPSHRDLDRN 361
QY 150 RTMSYRGTFQFCRRFPMNDKLGYSWFHVEHQRDMSNVFLFKYEDMHRDLVTWVEQ 209
DB 362 RPMGLHGTRQE---EMIDHRLTDREWABE---WKH-----LDHLLNCIMDMVEK 404
QY 210 LARFLGV--SCDKAQLEALTECHQLVDQCCNAEALPVGRG 248
DB 405 TRRSITVLRRCQEADEELNYWIRYS-----AEDLKGGG 441

RESULT 91
US-08-244-189-2
; Sequence 2, Application US/08244189
; Patent No. 5580727
; GENERAL INFORMATION:
; APPLICANT: Ohki, Misao
; APPLICANT: Kikuchi, Kimiko
; APPLICANT: Miyoshi, Hiroyuki
; APPLICANT: Kozu, Tomoko
; TITLE OF INVENTION: No. 5580727el DNAs, Polypeptides Encoded Thereby
; TITLE OF INVENTION: and Methods for Detecting the DNAs and Polypeptides
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/244,189
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 760-183P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 752 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-244-189-2

Query Match      5.8%; Score 88; DB 1; Length 752;
Best Local Similarity 21.7%; Pred. No. 1.1;
Matches 61; Conservative 38; Mismatches 92; Indels 90; Gaps 15;

QY 21 EFHGVRLPPFCRGKMEETANFVRPSDVWIVTPKSGTSLLOEVVLYSQGADPDEIGLM 80
DB 198 EFH-----SKLQEATNFPRLP---FVIFPKANLPLLORELLHCARLAKONPAOYL 245
QY 81 NIDEQL-----PV-----LEYQPGLDIIKELTSPRIK-----SHLP 113
DB 246 AQHEQLLDASTTSPVDSSELLDVNENG-----KRTPTDKENGDFREPLHSEHPKRP 301
QY 114 YRFLPSDLHNGDSKVIYARN-----PK-----DLVVS-----YQFHRSL----- 149
DB 302 CTISPGQKRYSPNGLSYQPNGLPHPTPPPPQHYRLDDMAIAHHYRDSYRHPSHRDLDRN 361
QY 150 RTMSYRGTFQFCRRFPMNDKLGYSWFHVEHQRDMSNVFLFKYEDMHRDLVTWVEQ 209
DB 362 RPMGLHGTRQE---EMIDHRLTDREWABE---WKH-----LDHLLNCIMDMVEK 404
QY 210 LARFLGV--SCDKAQLEALTECHQLVDQCCNAEALPVGRG 248
DB 405 TRRSITVLRRCQEADEELNYWIRYS-----AEDLKGGG 505
```

Db 312 EFH-----SKQEAENFPLRP---FVIFPLKANLPLLORELLHCARLAKQNPAYL 359  
QY 81 NIDEQL-----PV-----LEYDQGLDIKELTSPRLIK-----SHLP 113  
Db 360 AQHEQLLDASTTSPVDSESLLDVNEG-----KRTPTDRTKENGFDREPLHSEHPKRP 415  
QY 114 YRFLPSDLHGDSKVIYMARN-----PK-----DLVVS-----YQFHRSL----- 149  
Db 416 CTISPGQRYSPNNGUSYQNGLPHPTPPPPQHYRLDDMAIAHHYRDSYRHPSHRDLDRN 475  
QY 150 RTMSYRGTFQEFRCRPMNDKLGYSWFHVFQEFWEHRMDSNVFLFKYEDMHRDLVTMV 207  
Db 476 RPYGLHGTQE---EMIDHRLTDREWABE-----WKH-----LDHLLNCIMDMVEK 518  
QY 210 LARFLGV---SCDKAQLEALTECHQLVDOCCNAELPVG 248  
Db 519 TRRSUTVLRRCCQEADEELNYWIRYSYD-----AEDLKGGG 555

RESULT 92

US-09-345-473E-12  
; Sequence 12, Application US/09345473E  
; Patent No. 6558903  
; GENERAL INFORMATION:  
; APPLICANT: Hodge, Martin  
; TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof  
; FILE REFERENCE: 35800/183781  
; CURRENT APPLICATION NUMBER: US/09/345,473E  
; PRIOR FILING DATE: 1999-06-30  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 172  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(172)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-345-473E-12

Query Match 5.7%; Score 87.5; DB 2; Length 172;  
Best Local Similarity 28.8%; Pred. No. 0.14; Mismatches 35; Indels 23; Gaps 5;

Matches 30; Conservative 16; Mismatches 35; Indels 23; Gaps 5;  
QY 153 SYRGTFQBFQCR-----RFMNDKLGYSWFHVFQEFWEHRMDSNVFLFKYEDMHRDLVTMV 207  
Db 38 SYRALISAFSLRTRLDLDDFTCEKIGSG-PFSEVFKV-RHRASQGVMAKXNTLSSNRANWL 95  
QY 208 EQ-----LARFLGVSCDKAQLEALTEHCH-----QLVD 235  
Db 96 KEVQLMNLRLSHPNILRFMGVCVHOGQLHALTEYINSGNLEQLLD 139

RESULT 93

US-09-862-027-12  
; Sequence 12, Application US/09862027  
; Patent No. 6858418  
; GENERAL INFORMATION:  
; APPLICANT: Hodge, Martin R.  
; TITLE OF INVENTION: No. 6858418el Kinases and Uses Thereof  
; FILE REFERENCE: 35800/234862  
; CURRENT APPLICATION NUMBER: US/09/862,027  
; PRIOR FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: US 09/345,473  
; PRIOR FILING DATE: 1999-06-30  
; NUMBER OF SEQ ID NOS: 82  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 172  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:

; NAME/KEY: VARIANT  
; LOCATION: (1)...(172)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-862-027-12

Query Match 5.7%; Score 87.5; DB 2; Length 172;  
Best Local Similarity 28.8%; Pred. No. 0.14; Mismatches 35; Indels 23; Gaps 5;

Matches 30; Conservative 16; Mismatches 35; Indels 23; Gaps 5;  
QY 153 SYRGTFQBFQCR-----RFMNDKLGYSWFHVFQEFWEHRMDSNVFLFKYEDMHRDLVTMV 207  
Db 38 SYRALISAFSLRTRLDLDDFTCEKIGSG-PFSEVFKV-RHRASQGVMAKXNTLSSNRANWL 95  
QY 208 EQ-----LARFLGVSCDKAQLEALTEHCH-----QLVD 235  
Db 96 KEVQLMNLRLSHPNILRFMGVCVHOGQLHALTEYINSGNLEQLLD 139

RESULT 94

US-08-946-329A-49  
; Sequence 49, Application US/08946329A  
; Patent No. 6057091  
; GENERAL INFORMATION:  
; APPLICANT: Beachy, Philip A.  
; APPLICANT: Porter, Jeffrey A.  
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
; NUMBER OF SEQUENCES: 109  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/946,329A  
; FILING DATE: 07-OCT-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/061,323  
; FILING DATE: 07-OCT-1996  
; APPLICATION NUMBER: 08/729,743  
; FILING DATE: 10-JUL-1996  
; APPLICATION NUMBER: 08/567,357  
; FILING DATE: 04-DEC-1995  
; APPLICATION NUMBER: 08/349,498  
; FILING DATE: 02-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haile, Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07265/140001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619/678-5070  
; TELEFAX: 619/678-5099  
; INFORMATION FOR SEQ ID NO: 49:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 154 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-946-329A-49

Query Match 5.7%; Score 87; DB 2; Length 154;  
Best Local Similarity 24.6%; Pred. No. 0.13;  
Matches 32; Conservative 22; Mismatches 48; Indels 28; Gaps 6;

QY 154 YRGTFQBFQCRPMNDKLGYSWFHVFQEFWEHRMDSNVFLFKYED-MHRDLVTWVQOLAR 212  
Db 11 YNGGFKRM-----DELAVDGWQALD-----KNGSQVTFIPVQVWLHRD-----PKQVAD 55

QY	213	FLGVSCDKAQLEALTEHCHOLVDQC-----CNAEALPVGRGVGLWKDIFTYSMNE	263
		: : : :         : : : :   : : :   : : :	
Db	56	FVEETLDNGETFSITEKHLVETOCSPVSEEDINANPVPAERNVIG---DCFYIAHRK	112

Qy 264 KFDLVYKQKM 273  
| : : |  
pb 113 KSOMYORVKV 122

RESULT 95  
US-09-562-914-49  
; Sequence 49, Application US/09562914  
; Patent No. 6911528  
; GENERAL INFORMATION:  
; APPLICANT: Beachy, Philip A.  
; Porter, Jeffrey A.  
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
; NUMBER OF SEQUENCES: 109  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037

Query Match 5.7%; Score 87; DB 2; Length 154;  
Best Local Similarity 24.6%; Pred. No. 0.13;  
Matches 32; Conservative 22; Mismatches 48; Indels 28; Gaps 6;

QY	213	FLGVSCDKAQLEALTEHCHQLVDQC-----CNAEALPVGRGVLWKDIFVTSMNE	263
Dp	56	FVEETLDNGCTESTETKHLVFTOCSPVSYSEDINANPVAERNVG---DCFYTAHRK	112

Qy 264 KFDLVTKQKM 273  
| : | :  
Db 113 KSQMYQRVKV 122

RESULT 96  
US-09-248-796A-14484  
; Sequence 14484, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:

	Query Match	5.7%; Score 86.5; DB 2; Length 499;
	Best Local Similarity	23.9%; Pred. No. 0.9;
	Matches	Matched 47; Conservative Matches 26; Mismatches 71; Indels 53; Gaps 10;
Qy	40 NFVVRPSDVM-----IVTPKSGTSL---LQEVLVYVGADPDEIGLMN- 81     :     :	:
Dd	167 NFDIHDSVWLSYSQSHTSHDFKYFGSDKSNGIIKTDLRSVFPTNTSNDDEFGSNY 226     :	: ::::
Qy	82 ----IDQLPVLEYPQGGLPIIKELTSPLRIKSHLPYR----FLPSDLHUNGDSKVYYMARN 134     :	:
Dd	227 ENNYIDEKLGI-----STLIAKSTSPILSIICHEPEQDTIFAT-----NMESMPTRYI--N 273     :	:
Qy	135 PKDLVVSYGYPHRLSRMTWSRGTFQECFRPMNDKLGVSFWHFHQEFWEHRMDSNLVLF 194     :     :	:
Dd	274 PDTCLSQYQLRCLOLYSYN-----TNNGGVGN--DNDELHLHDGGSNP---- 318     :	:
Qy	195 KYEDMRDLVTMWBOLA 211     :	:
Dd	319 PTELDNSGFYDLISHLS 335     :	:

```

RESULT 97
US-09-445-353E-2
; Sequence 2, Application US/09445353E
; Patent No. 6812336
; GENERAL INFORMATION:
; APPLICANT: Rosenfeld, Michael G.
; APPLICANT: Glass, Christopher K.
; APPLICANT: Rose, David W.
; APPLICANT: Torchia, Joseph
; TITLE OF INVENTION: A Transcription Factor Coactivator Protein, p/CIP
; FILE REFERENCE: 6627-PAL021
; CURRENT APPLICATION NUMBER: US/09/445,353E
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: PCT/US98/12263
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/049,452
; PRIOR FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1402
; TYPE: PRP
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature

```





Best Local Similarity 21.9%; Pred. No. 0.62; Matches 49; Conservative 44; Mismatches 69; Indels 62; Gaps 14;

Qy	51	VTPKSGTSLQEVVYLVSOQA-----DPDEI-----GLMNIDEQLPVLEY	91
Db	13	IIYPVADSQWYSHVHYVPKGGMTVVPNDKDELIPQIVTGYRMVSDFRKLNKATKKDHY	72
Qy	92	POGLDIKELTSPRLIKSHLPYRPLPSDLHNGDSKVIYMARNPKDLVVSYQFHRSLRT	151
Db	73	PLPFIDQMLE-----RLSK-HTHFFFL--DGYSGFPSQ-IPVSKGDOEKT-----FTCTFGT	120
Qy	152	MSYR-----GTFO-----BFCRR-----FMNDKLGYSWPE-----HVQEFWEH	185
Db	121	FAYRRMPGLCNAPATFORCMAIFSDFCFKIVEVFMDDFSVYGTSTFDDCLSLNLDRLQR	180
Qy	186	RMDSNVFLKYEDMH---RDLVTMVEQLARFLGVSCDKAQLEAL	226
Db	181	CEDTN-LVLNCEKCHFMVNEGIVLGHKISE-IGIEVDKAKVDAI	222

Search completed: May 17, 2006, 11:06:27  
Job time : 56 secs

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 17, 2006, 09:15:13 ; Search time 75.9 Seconds  
(without alignments)  
4371.567 Million cell updates/sec

Title: US-10-768-158-1  
Perfect score: 4401  
Sequence: 1 9cgcggcgagcggcgccg.....aaaaaaaaaaaaaaaaaaaaa 2419

Scoring table:  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 5179358

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/abss/ABSSWEB.spool/US10768158/runat.16052006.163420.5271/app.query.fasta.1  
-DB=A Geneseq -QFMT=Eastan -SURFIX=n2p.rag -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=1000 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02p  
-USER=US10768158 @CGN.1.1.348 @runat.16052006.163420.5271 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOC=100 -LONGLOG -DEV TIMEOUT=120  
-WAEN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A\_Geneseq.8.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*  
9: Geneseqp2005s.\*  
10: Geneseqp2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1530	34.8	284	3	Aay79215 Human tra
2	1530	34.8	284	5	Aau78268 Human ner
3	1530	34.8	284	6	Abro1809 Human can
4	1530	34.8	284	8	Adq89158 Human uro
5	1530	34.8	284	8	Adr44884 Polypepti
6	1530	34.8	284	9	Adz21933 Human sul
7	1506	34.2	284	6	Abm04807 Rat sulfo
8	1142.5	26.0	218	6	Abro1860 Human can

171	6	ABR01861	171	9	ABR01861	Human can
171	9	ADZ21933	171	9	ADZ21933	Human SUL
159	5	AAU78268	159	5	AAU78268	Rat nervo
105	4	AAU18674	105	4	AAU18674	Renal and
105	6	ABU97289	105	6	ABU97289	Human pol
296	5	ADI17131	296	5	ADI17131	Murine NO
102	4	ABBI6232	102	4	ABBI6232	Human ner
1305	4	ABG27554	1305	4	ABG27554	Novel hum
295	2	AAW23657	295	2	AAW23657	E6AP-bind
295	7	ADD18680	295	7	ADD18680	Human dis
295	8	ADP24302	295	8	ADP24302	PRO polyp
295	9	ADW71803	295	9	ADW71803	Human phe
296	5	ADI17130	296	5	ADI17130	Rat NOVX
285	8	ADY80433	285	8	ADY80433	Rat aryl
291	7	ADD45448	291	7	ADD45448	Rat Prote
291	7	ADE57139	291	7	ADE57139	Rat Prote
291	7	ADE57135	291	7	ADE57135	Rat Prote
291	7	ADE57147	291	7	ADE57147	Rat Prote
291	7	ADE57143	291	7	ADE57143	Rat Prote
295	3	AAI67294	295	3	AAI67294	Human STP
295	7	ADP65305	295	7	ADP65305	Human sul
295	8	ADP24022	295	8	ADP24022	PRO polyp
295	9	ADW71800	295	9	ADW71800	Pig pheno
295	9	ADW71802	295	9	ADW71802	Human phe
295	9	ADW71805	295	9	ADW71805	Pig pheno
295	9	ADW71815	295	9	ADW71815	Pig pheno
295	7	ADD18776	295	7	ADD18776	Human dis
295	7	ADE57149	295	7	ADE57149	Human Pro
295	7	ADE57137	295	7	ADE57137	Human Pro
295	7	ADE57141	295	7	ADE57141	Human Pro
295	7	ADE57145	295	7	ADE57145	Human Pro
295	7	ADI15901	295	7	ADI15901	Human PP
295	8	ADP23588	295	8	ADP23588	PRO polyp
295	8	ADR97362	295	8	ADR97362	Human SUL
295	9	ADW71801	295	9	ADW71801	Human phe
302	5	ADI17050	302	5	ADI17050	Human NOV
302	7	ADF76915	302	7	ADF76915	Novel hum
302	9	ADZ70333	302	9	ADZ70333	Human pro
296	5	ADI17132	296	5	ADI17132	Rat NOVX
327	8	ABM84002	327	8	ABM84002	Human dia
327	8	ABM84001	327	8	ABM84001	Human dia
327	8	ABM82615	327	8	ABM82615	Human dia
327	8	ABM84003	327	8	ABM84003	Human dia
327	8	ABM84004	327	8	ABM84004	Human dia
269	5	ADI17319	269	5	ADI17319	Polypepti
269	5	ADI17331	269	5	ADI17331	Polypepti
310	7	ADD12573	310	7	ADD12573	Human ENZ
296	6	ADA10983	296	6	ADA10983	Human CDN
294	2	AAW40498	294	2	AAW40498	Human EST
294	2	AAW44247	294	2	AAW44247	Human oes
294	7	ADD48896	294	7	ADD48896	Human Pro
307	5	ADI17049	307	5	ADI17049	Chicken N
304	4	AAE12209	304	4	AAE12209	Human ST
304	5	ABB81792	304	5	ABB81792	Human sul
295	5	ABB81793	295	5	ABB81793	Rat Prote
304	5	ABB81794	304	5	ABB81794	Human sul
283	5	ABB81794	283	5	ABB81794	Human sul
304	5	AAM50843	304	5	AAM50843	Orphan G
304	5	ADI17048	304	5	ADI17048	Murine NO
304	4	AAE12210	304	4	AAE12210	Human ST
316	4	ABB66006	316	4	ABB66006	Drosophil
304	4	AAE12208	304	4	AAE12208	Human ST
304	5	ADI17046	304	5	ADI17046	Human NOV
305	6	ABU12099	305	6	ABU12099	Novel hum
304	4	AAE05178	304	4	AAE05178	Human dru
261	7	ADD12583	261	7	ADD12583	Human ENZ
295	5	ADI16670	295	5	ADI16670	Human NOV
295	8	ADN42324	295	8	ADN42324	Human nov
305	5	AAG68266	305	5	AAG68266	Human POL
305	7	AAE38814	305	7	AAE38814	Human sul
297	5	ADR19665	297	5	ADR19665	Human dru
262	8	ABM84008	262	8	ABM84008	Human dia

82 384.5 8.7 262 8 ABM82616  
 83 384.5 8.7 262 8 ABM84005  
 84 383 8.7 253 8 ABM82696  
 85 383 8.7 324 5 ADI16668  
 86 383 8.7 324 8 ADN42322  
 87 382 8.7 76 4 AAU18710  
 88 382 8.7 76 4 ABB14904  
 89 382 8.7 76 6 ABU97325  
 90 376.5 8.6 254 8 ABM84007  
 91 376.5 8.6 254 8 ABM84009  
 92 376.5 8.6 254 8 ABM84006  
 93 376.5 8.6 254 8 ABM82619  
 94 376.5 8.6 254 8 ABM82617  
 95 376.5 8.6 254 8 ABM82618  
 96 376.5 8.6 254 8 ABM82620  
 97 363.5 8.3 259 7 ADG76311  
 98 363.5 8.3 338 4 ABB61517  
 99 359.5 8.2 346 4 ABB62612  
 100 357.5 8.1 254 8 ABM82695

## ALIGNMENTS

## RESULT 1

ID AAY79215 standard; protein; 284 AA.

AC AAY79215;

XX 19-JUN-2000 (first entry)

DT Human transferase TRNSFS-7.

DE Transferase; TRNSFS-7; human; antitumour; cancer;

KW gastrointestinal disorder; developmental disorder; genetic disorder;

KW neurological disorder; reproductive disorder; smooth muscle disorder;

KW immunological disorder; inflammation; diagnosis; therapy;

KW sulfotransferase.

XX Homo sapiens.

XX Location/Qualifiers

FT Modified-site 4 /note= "potential O-phosphorylation"

FT Modified-site 11 /note= "potential O-phosphorylation"

FT Domain 23..272 /note= "potential O-phosphorylation"

FT Modified-site 104 /note= "sulfotransferase protein domain"

FT Modified-site 148 /note= "potential O-phosphorylation"

FT Modified-site 153 /note= "potential O-phosphorylation"

FT Modified-site 157 /note= "potential O-phosphorylation"

FT Modified-site 172 /note= "potential O-phosphorylation"

FT Modified-site 174 /note= "potential O-phosphorylation"

FT Modified-site 205 /note= "potential O-phosphorylation"

FT Modified-site 260 /note= "potential O-phosphorylation"

FT Modified-site /note= "potential O-phosphorylation"

XX WO200014251-A2.

XX 16-MAR-2000.

XX 09-SEP-1999; 99WO-US020989.

XX 10-SEP-1998; 98US-00150657.

XX 04-NOV-1998; 98US-00186779.

ABM82616 Human dia  
 ABM84005 Human dia  
 ABM82696 Human dia  
 ADI16668 Human NOV  
 ADN42322 Human nov  
 AAU18710 Renal and  
 ABB14904 Human ner  
 ABU97325 Human pol  
 ABM84007 Human dia  
 ABM84009 Human dia  
 ABM84006 Human dia  
 ABM82619 Human dia  
 ABM82617 Human dia  
 ABM82618 Human dia  
 ABM82620 Human dia  
 ADG76311 Human Inc  
 ABB61517 Drosophil  
 ABB62612 Drosophil  
 ABM82695 Human dia

PR 11-MAY-1999; 99US-0133642P.

XX (INCY-) INCYTE PHARM INC.

XX Tang YT, Corley NC, Guegler KJ, Baughn MR, Lal P, Yue H;

XX Hillman JU, Azimzai Y;

XX WPI; 2000-256996/22.

DR N-PSDB; AA294207.

XX Human transferase proteins useful for preventing, diagnosing and treating  
 cancers and developmental, gastrointestinal, genetic, immunological,  
 neurological, reproductive and smooth muscle disorders.

XX Claim 1; Page 85-86; 113pp; English.

XX The present sequence is that of human transferase TRNSFS-7, 1 of 15  
 claimed novel human transferase proteins of the invention (see AAY79209-  
 23). The sequence was deduced from a cDNA clone (see AA294207) isolated  
 from a cerebellum library. It shows homology to sulfotransferase  
 proteins. TRNSFS-7 is expressed in nervous, endocrine and development  
 tissues, especially those associated with neurological disorders,  
 inflammation and cancer. The new human transferases and polynucleotides  
 can be used in the diagnosis, prevention and treatment of cancer,  
 developmental disorders, gastrointestinal disorders, genetic disorders,  
 immunological disorders, neurological disorders, reproductive disorders,  
 and smooth muscle disorders. The polypeptides can also be used to raise  
 antibodies, and to screen for agonists and antagonists of transferase  
 activity

XX SQ Sequence 284 AA;

## Alignment Scores:

Pred. No.: 3.01e-136 Length: 284  
 Score: 1530.00 Matches: 284  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 34.8% Indels: 0  
 DB: 3 Gaps: 0

US-10-768-158-1 (1-2419) x AAY79215 (1-284)

QY 21 ATGGCGGAGCGAGCGGAGAGCCCGGAGGAGTTTCGAGAGCAAGTACTTC 80  
 DB 1 MetAlaGluSerGluAlaGluThrProSerThrProGluPheGluSerLeuTyrPhe 20  
 QY 81 GAGTTCATGCGTGGCGCTGCCCGCTTCGCGCGGAGAGATGGAGAGATCGCCAAC 140  
 DB 21 GluPheHisGlyValArgLeuProProPheCysArgGlyLysMetGluGluLeuAlaAsn 40  
 QY 141 TTCCCGGTGCGGCGGAGCGAGTGGATCGTACCTACCCAGTCCGCGCAGCTTG 200  
 DB 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60  
 QY 201 CTGCAGGAGGTGGTCTACTTGGTGAGCGGCGCTGACCCCGATGAGATCGGTTGATG 260  
 DB 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80  
 QY 261 AACATCGACGACGAGTCCCGTCTCTGGAGTACCCACAGCGCGGCTGGACATCATCAAG 320  
 DB 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100  
 QY 321 GAATGACCTCTCCCGCTCATCAAGGCGACCTGCCCTTACCGCTTCTCCCTCTGAC 380  
 DB 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120  
 QY 381 CTCCACAATGGAGACTCCAAGGTCTATATGGTCTGCAACCCCAAGATCTGGTGTG 440  
 DB 121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuValVal 140  
 QY 441 TCTTATTATCATGTTCCACCGCTCTCTCGGACCATGAGTACCGAGGACCTTTCAAGAA 500  
 DB 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160

QY 501 TTCTGCCGAGGTTTATGAATGATAAGCTGGCTACCGCTCCTCGTTTGACACGTCGAG 560  
 |||||  
 DB 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180  
 |||||  
 QY 561 GAGTTCTGGGAGCACCGCATCGACTCGAAGCTGCTTTTCTCAAGTATGAGACATGCAT 620  
 |||||  
 DB 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHis 200  
 |||||  
 QY 621 CGGACCTGCTGACGATGGTGAGCAGCTGGCCAGATTCTCTGGGGGTCTCTGTGACAAG 680  
 |||||  
 DB 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220  
 |||||  
 QY 681 GCCAGCTGGAAGCCCTGACCGAGACACTGCCACAGCTGGTGGACAGTGTGCAACGCT 740  
 |||||  
 DB 221 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240  
 |||||  
 QY 741 GAGCCCTGCCCGTGGCGGGGAAGAGTGTGGGTGTGGGAAGGACATCTTCACCGTCTCC 800  
 |||||  
 DB 241 GluAlaLeuProValGlyArgGlyValGlyLeuTrpLysAspIlePheThrValSer 260  
 |||||  
 QY 801 ATGAATGAGAAGTTTGACTGGTGTATATAACAGAAAGATGGAAAGTGTGACCTCACGTTT 860  
 |||||  
 DB 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280  
 |||||  
 QY 861 GACTTTTATTTA 872  
 |||||  
 DB 281 AspPheTyrLeu 284  
 |||||

RESULT 2  
 AAU78268  
 ID AAU78268 standard; protein; 284 AA.  
 AC AAU78268;  
 DT  
 XX 05-JUN-2002 (first entry)  
 XX Human nervous system-specific sulfotransferase, SULTn, protein sequence.  
 XX  
 KW Human; nervous system-specific sulfotransferase; SULTn; sulfonation;  
 KW neurological disorder; multiple sclerosis; Huntington's chorea;  
 KW nervous system-mediated muscular atrophy; amyotrophic lateral sclerosis;  
 KW myasthenia gravis; Alzheimer's disease; senile dementia; Schizophrenia;  
 KW bipolar disorder; cytosolic sulfotransferase gene family;  
 KW cerebellar ataxia; chromosome 22q13.  
 XX  
 OS Homo sapiens.  
 XX  
 PH Key Location/Qualifiers  
 FT Binding-site 54..61 /label= PAPS\_5'\_phosphosulfate\_binding\_site  
 FT Binding-site 129..144 /label= PAPS\_3'\_phosphate\_binding\_site  
 FT Binding-site 246..258 /label= PAPS\_3'\_phosphate\_binding\_site  
 XX  
 PN W0200218541-A2.  
 XX  
 XX 07-MAR-2002.  
 XX  
 XX 04-SEP-2001; 2001WO-US041986.  
 XX  
 XX 01-SEP-2000; 2000US-0229929P.  
 XX  
 XX (UYBO-) UNIV BOSTON.  
 XX  
 XX Farb DH, Martin S;  
 PI  
 XX WPI; 2002-281063/32.  
 DR N-PSDB; ABK12289.  
 XX  
 XX New nervous system-specific sulfotransferase (SULTn) polypeptide for  
 PT treating neurological disorders including multiple sclerosis, nervous

PT system-mediated muscular atrophy and Huntington's chorea.  
 XX Claim 2; Fig 1a; 46pp; English.  
 PS  
 XX The present invention relates to a new nervous system-specific  
 CC sulfotransferase (SULTn) polypeptide. The SULTn polypeptide of the  
 CC invention comprises a nervous system-specific sulfotransferase,  
 CC comprising a sequence of 284 amino acids, given in the specification. The  
 CC invention is useful for identifying an effector of nervous system-  
 CC specific sulfonation by contacting the polypeptide with a candidate  
 CC effector compound and assaying the ability of the candidate effector  
 CC compound to effect the sulfotransferase function of the polypeptide. The  
 CC invention is also useful for inhibiting nervous system-specific  
 CC sulfonation by contacting the polypeptide under physiological conditions  
 CC with a compound capable of interacting with and inhibiting the  
 CC sulfonation activity of the polypeptide, where the compound is an  
 CC antibody, peptide, polypeptide, nucleic acid, organic molecule or an  
 CC inorganic molecule. The polypeptide of the invention is also useful for  
 CC treating a neurological disorder, including multiple lateral sclerosis,  
 CC system-mediated muscular atrophy, amyotrophic lateral sclerosis,  
 CC Huntington's chorea, myasthenia gravis, Alzheimer's disease, senile  
 CC dementia, Schizophrenia, bipolar disorder and cerebellar ataxia. The  
 CC present amino acid sequence represents the human SULTn protein of the  
 CC invention. This sequence is encoded by the human SULTn gene located on  
 CC chromosome 22q13. SULTn is a member of the cytosolic sulfotransferase  
 CC gene family  
 XX  
 SQ Sequence 284 AA;

Alignment Scores:  
 Pred. No.: 3.01e-136 Length: 284  
 Score: 1530.00 Matches: 284  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 34.8% Indels: 0  
 DB: 5 Gaps: 0

US-10-768-158-1 (1-2419) x AAU78268 (1-284)

QY 21 ATGCGGAGAGCGAGGCGGAGACCCCGAGACCCCGGGGAGTTCGAGAGCAAGTACTTC 80  
 |||||  
 DB 1 MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerLysTyrPhe 20  
 |||||  
 QY 81 GAGTTCATGGCGTGGCGGTGCGCGCTTCTGCGGGGAAGATGGAGGAGATCGCCAAC 140  
 |||||  
 DB 21 GluPheHisGlyValArgLeuProProPheCysArgGlyLysMetGluGluIleAlaAsn 40  
 |||||  
 QY 141 TTCGCGGTGGCGCCGAGCGAGCTGTGATCGTCACTACCCAGTCCCGCACCAGCTTG 200  
 |||||  
 DB 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60  
 |||||  
 QY 201 CTGCGAGAGGTGGTCTACTTTGGTGAGCCAGGGCGCTGACCCCGATGAGATCGGCTTGATG 260  
 |||||  
 DB 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80  
 |||||  
 QY 261 AACATGACGAGCAGCTCCCGGTCTGAGTACCCACAGCCGGCGCTGGACATCATCAAG 320  
 |||||  
 DB 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100  
 |||||  
 QY 321 GAACTGACCTCTCCCGCTCATCAGAGCCACCTGCCCTACCGCTTCTCGCCTCTGAC 380  
 |||||  
 DB 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120  
 |||||  
 QY 381 CTCCACAATGGAGACTCCAAGGTCACTATATGGCTCGCAACCCCAAGGATCTGGTGGTG 440  
 |||||  
 DB 121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuValVal 140  
 |||||  
 QY 441 TCTTATTATCAGTTCACCGCTCTCTGCGGACCATGAGCTACCGAGCACCTTTCAAGAA 500  
 |||||  
 DB 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160  
 |||||  
 QY 501 TTCTGCCGAGGTTTATGATGATAGCTGGGTACGGCTCTCTGGTTTGACGCTGCAG 560  
 |||||

Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180  
QY 561 GAGTCTCTGGGAGCACCAGCTCGAAGCTGCTTTTCTCAAGTATGAAGACATGCAT 620  
Db 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHis 200  
QY 621 CGGACCTGTGTGACAGTGTGGAGCAGCTGCGCCAGATTCTTGGGGGTGTCTGTGACAAG 680  
Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220  
QY 681 GCCAGCTGGAAGCCCTGACGAGACATGCCACAGCTGTGTGGACCAAGTGTGCAACGCT 740  
Db 221 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240  
QY 741 GAGCCCTGCGCGGCGCGGAGAGTGGCTGTGGAGGACATCTTCACCGTCTCC 800  
Db 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 260  
QY 801 ATGAATGAGAAGTTTGACTTGTGTATATAAAGAGATGGAAAGTGTGACCTCACGTTT 860  
Db 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280  
QY 861 GACTTTTATTTA 872  
Db 281 AspPheTyrLeu 284

RESULT 3  
ABR01809  
ID ABR01809 standard; protein; 284 AA.  
XX AC ABR01809;  
XX XX  
DT 19-MAY-2003 (first entry)  
XX  
DE Human cancer-related protein, 192P2G7.  
XX  
KW Human; cytostatic; vaccine; cancer; immune response.  
XX  
OS Homo sapiens.  
XX  
PN WO200283921-A2.  
XX  
PD 24-OCT-2002.  
XX  
PF 10-APR-2002; 2002WO-US011654.  
XX  
PR 10-APR-2001; 2001US-0282739P.  
XX  
PR 10-APR-2001; 2001US-0283112P.  
XX  
PR 25-APR-2001; 2001US-0286630P.  
XX  
XX (AGEN-) AGENSYS INC.  
XX  
XX Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;  
PI Morrison K, Morrison RK, Raitano AB;  
XX  
XX WPI; 2003-075555/07.  
DR N-PSDB; ABZ78140.  
XX  
XX New composition comprising a substance that modulates the structure of  
PT proteins and polynucleotides, useful for therapeutic, prognostic and  
PT diagnostic reagents for eliciting cellular or humoral immune response in  
PT cancer patients.  
XX  
XX Claim 12; Fig 2R; 1021pp; English.  
XX  
XX The present invention relates to novel human cancer-related genes and  
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and  
CC proteins are useful for eliciting a humoral or cellular immune response.  
CC The genes are useful as probes and primers for the amplification and/or  
CC detection of genes, mRNAs or their fragments, as reagents for the  
CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
CC directing the expression of the protein, as tools for modulating or  
CC inhibiting the expression of genes and/or translation of transcripts, and

CC as therapeutic agents. The proteins and peptides are useful as  
CC therapeutic, prognostic and diagnostic reagents for cancer  
XX  
SQ Sequence 284 AA;  
Alignment Scores:  
Pred. No.: 3,01e-136 Length: 284  
Score: 1530.00 Matches: 284  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 34.8% Indels: 0  
DB: 6 Gaps: 0  
US-10-768-158-1 (1-2419) x ABR01809 (1-284)  
QY 21 ATGCGGAGAGCAGGCGGAGACCCCGAGACCCCGGGGGAGTTCGAGAGCAAGTACTTC 80  
Db 1 MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerLysTyrPhe 20  
QY 81 GAGTTCATGCGTGGCGCTGCCCGCTTCTGCCGGGGAAGATGGAGGAGATCGCCAAC 140  
Db 21 GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluGluLeuAlaAsn 40  
QY 141 TTCCCGTGGCGGCCAGCGACGTGTGGATCGTCACTTACCTACCCCAAGTCCGGCACCAGCTTG 200  
Db 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60  
QY 201 CTGACGAGGTGTCTACTTGTGTGAGCCAGGGCGGTGACCCCGATGAGATCGGCTTGATG 260  
Db 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80  
QY 261 AACATGACGAGCAGCTCCCGTCTGGAGTACCCACAGCCGGCGCTGGACATCATCAAG 320  
Db 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleLeuLys 100  
QY 321 GAATGACCTCTCCCGCTTCATCAAGACGACCTGCCCTTACCGCTTTCTGCCCTCTGAC 380  
Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120  
QY 381 CTCACATGAGACTCCAAAGTCACTATATGGCTCGGAACCCCAAGGATCTGGTGGTG 440  
Db 121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuValVal 140  
QY 441 TCTTATTATCAGTTCCACCGCTCTCTGCGGACCATGAGTACCGAGGACCTTTTCAAGAA 500  
Db 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160  
QY 501 TTCTGCGGAGGTTTATGAATGAAGCTGGCTACGGCTCCTCGTGTGTGAGCAGCTCAG 560  
Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180  
QY 561 GAGTTCGGGAGCACCAGCATGGACTCGAACGTCTTTTCTCAAGTATGAAGACATGCAT 620  
Db 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHis 200  
QY 621 CGGACCTGTGTGACAGTGTGGAGCAGCTGCGCCAGATTCTTGGGGGTGTCTGTGACAAG 680  
Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220  
QY 681 GCCAGCTGGAAGCCCTGACGAGCAGCTGCCACAGCTGTGTGGACCAAGTGTGCAACGCT 740  
Db 221 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240  
QY 741 GAGCCCTGCGCGGCGCGGAGAGTGGCTGTGGAGGACATCTTCACCGTCTCC 800  
Db 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 260  
QY 801 ATGAATGAGAAGTTTGACTTGTGTATATAAAGAGATGGAAAGTGTGACCTCACGTTT 860  
Db 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280  
QY 861 GACTTTTATTTA 872  
XXXXXXXXXXXXXXXXXXXX







PT member 1.  
 XX Claim 15; SEQ ID NO 1; 83pp; English.  
 XX

The invention describes a method of diagnosing or prognosticating a neurodegenerative disease in a subject, or determining whether a subject is at increased risk of developing the disease. The method comprises determining a level and/or an activity of a transcription product and/or a translation product of a gene coding for a cytosolic sulfotransferase family 4A member 1 and/or a fragment, or derivative or variant of the transcription or translation product, in a sample obtained from the subject and comparing the level and/or the activity to a reference value representing a known disease or health status, therefore diagnosing or prognosticating the neurodegenerative disease in the subject, or determining whether the subject is at increased risk of developing the neurodegenerative disease. Protein molecules comprising fully defined 284 given in the specification, the protein molecules being translation products of the gene coding for a cytosolic sulfotransferase family 4A member 1, or its fragments, or derivatives, or variants, are useful as diagnostic targets for detecting a neurodegenerative disease, preferably Alzheimer's disease or as screening targets for reagents or compounds preventing, or treating, or ameliorating the disease. An antibody specifically immunoreactive with an immunogen, where the immunogen is a translation product of a gene coding for a cytosolic sulfotransferase family 4A member 1, SEQ ID NO. 1 or 2, or its fragment, or derivative, or variant is useful for detecting the pathological state of a cell in a sample obtained from a subject, comprising immunocytochemical staining of the cell with the antibody, where an altered degree of staining, or an altered staining pattern in the cell compared to a cell representing a known health status indicates a pathological state of the cell which relates to a neurodegenerative disease, preferably to Alzheimer's disease. The genetically altered non-human animal is useful for screening, testing, and validating compounds, agents, and modulators in the development of diagnostics and therapeutics to treat neurodegenerative diseases, in particular Alzheimer's disease. An agent(s) which directly or indirectly affect an activity and/or a level of a gene coding for a cytosolic sulfotransferase family 4A member 1, and/or a transcription product or a translation product of a gene coding for a cytosolic sulfotransferase family 4A member 1, and/or their fragment, or derivative, or variant is useful for treating or preventing a neurodegenerative disease, in particular Alzheimer's disease. This is the amino acid sequence of human sulfotransferase family 4A member 1 splice variant 1 (SULT4A1sv1).

XX  
 SQ Sequence 284 AA;

Alignment Scores:  
 Pred. No.: 3.01e-136 Length: 284  
 Score: 1530.00 Matches: 284  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 34.8% Indels: 0  
 DB: 9 Gaps: 0

US-10-768-158-1 (1-2419) x ADZ21933 (1-284)

QY 21 ATGCGGAGAGCGAGCGGAGACCCCGGAGGAGTTTCGAGAGCAAGTACTTC 80  
 DB 1 MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerLysTyPhe 20  
 QY 81 GAGTTCCATGCGGTGCGGCTGCCGCCCTTCTGCGCGGGAGAGATGGAGGATCGCAAC 140  
 DB 21 GluPheHisGlyValArgLeuProProPheCysArgGlyLysMetGluGluAlaAsn 40  
 QY 141 TTCCCGGTGGCGGCCAGGAGCTGTGATCTGACCTACCCCAAGTCCGGCAGCTTG 200  
 DB 41 PheProValArgProSerAspValTrpIleValThrTyPProLysSerGlyThrSerLeu 60  
 QY 201 CTGCAGGAGGTGGTCTACTTGGTGCAGCGGCGCTGACCCCGATGATCGGCTTGATG 260  
 DB 61 LeuGlnGluValValTyLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80

QY 261 AACATCAGCAGCAGCTCCCGGTCTCGAGTACCCACAGCCGGGCTTGACATCATCAAG 320  
 DB 81 AsnIleAspGluGlnLeuProValLeuGluTyPProGlnProGlyLeuAspIleLeLys 100  
 QY 321 GAACTGACCTCTCCCGCCTCATCAGAGCCACCTGCCCTACCGCTTTCTGCCCTCTGAC 380  
 DB 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProLysArgPheLeuProSerAsp 120  
 QY 381 CTCACAAATGAGACTCCAAAGGTCACTATATATGCTCGCAACCCCAAGGATCTCGTGGTG 440  
 DB 121 LeuHisAsnGlyAspSerLysValIleTyPMetAlaArgAsnProLysAspLeuValVal 140  
 QY 441 TCTATTATCAGTTCACCCCTCTCTCGCGACCATGAGCTACCGAGGACCTTTTCACAGAA 500  
 DB 141 SerTyTyPTrpGlnPheHisArgSerLeuArgThrMetSerTyPArgGlyThrPheGlnGlu 160  
 QY 501 TTCTGCCGAGGTTTATGAATGATGAAGCTGGGCTACGGCTCCTGGTTTGGACAGCTGCAG 560  
 DB 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyPArgSerTrpPheGluHisValGln 180  
 QY 561 GAGTTCTGGAGCACCACATGGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATGCAT 620  
 DB 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyPArgMetHis 200  
 QY 621 CGGACCTGTGAGCAGTGGTGGAGCAGCTGGCCAGATTCCTGGGGTGTCTCTGTGACAAG 680  
 DB 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220  
 QY 681 GCCCAGCTGGAAGCCCTGACGGAGCACTGCCACAGCTGGTGGACCAAGTGTCTGCAACGCT 740  
 DB 221 AlGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240  
 QY 741 GAGGCCCTGCGGCGGGCGGGAAGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCC 800  
 DB 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 260  
 QY 801 ATGAATGAGAAGTTGACTTGGTGTATAAACAGAGATGGAAAGTGTGACCTCAGCTTT 860  
 DB 261 MetAsnGluLysPheAspLeuValTyPLeuLysMetGlyLysCysAspLeuThrPhe 280  
 QY 861 GACTTTTATTATA 872  
 DB 281 AspPheTyPLeu 284

RESULT 7  
 ABM04807  
 ID ABM04807 standard; protein; 284 AA.  
 XX AC ABM04807;  
 XX DT 22-SEP-2003 (first entry)  
 XX DE Rat sulfotransferase-like protein.  
 XX KW spinal cord; neuropathic pain; central sensitisation pain; pain;  
 XX KW analgesic; gene therapy.  
 XX OS Rattus norvegicus.  
 XX PN EP1284298-A2.  
 XX PD 19-FEB-2003.  
 XX PF 26-JUL-2002; 2002EP-00255229.  
 XX PR 27-JUL-2001; 2001GB-00018354.  
 XX PR 07-FEB-2002; 2002GB-00002883.  
 XX PA (WARN ) WARNER LAMBERT CO.  
 XX PI Brooksbank RA, Dixon AK, Lee K, Pinnock RD;  
 XX WPI; 2003-543489/52.



Query Match:	26.0%	Indels:	45
DB:	6	Gaps:	1
US-10-768-158-1 (1-2419) x ABR01860 (1-218)			
QY	88	ATGGCGTGGCGCTGCGCCCTCTCTCGCGGGAAGATGGAGAGATCGCCAACTTCCCGG	147
Db	1	MetAlaCysGlyCysArgProSerAlaAlaGlyArgSerProThrSerArg	20
QY	148	TGCGGCCCGAGCGCTGTGATCGTCACTACCCCAAGTCGGCACCAGCTTGCTGCAGG	207
Db	21	CysGlyProAlaThrCysGlySerSerProThrProSerPro-----	34
QY	208	AGGTGCTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCG	267
Db	34	-----	34
QY	268	ACGAGCAGCTCCCGTCTCGAGTACCCACAGCGGCGCTGGACATCATCAAGGAACCTGA	327
Db	35	-----Glu--LeuT	37
QY	328	CCTCTCCCGCTCATCAAGAGCACCTGCCCTACCGCTTCTGCGCTCTGACCTCCACA	387
Db	37	hrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAspLeuHisA	57
QY	388	ATGGAGACTCCAAGGTCATCTATATGGCTCGCAACCCCAAGGATCTGGTGGTCTTATT	447
Db	57	enGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuValSerTyrT	77
QY	448	ATCAGTTCACCGCTCTCGGACCATGAGTACCCAGGACCTTTCAGAACTTCGCC	507
Db	77	yrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGluPheCysA	97
QY	508	GGAGGTTTATGAATGATAAGCTGGGCTACGCTCTCTGTTTGAGCACGTGCAGGAGTTCT	567
Db	97	rgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGlnGluPheT	117
QY	568	GGAGCACCCGATGGACTCGAACGTGCTTTTCTCAAAGTATGAAGACATCATCGGACC	627
Db	117	rpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHisArgAspL	137
QY	628	TGGTGACGATGGTGAGCAGCTGGCCAGATCTCTGGGGGTGCTCTGTGACAAGGCCAGC	687
Db	137	euValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLysAlaGlnL	157
QY	688	TGGAAGCCCTGACGGAGCACTGCCACAGCTGGTGACCACTGTCTGCAACGCTGAGGCC	747
Db	157	euGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAlaGluAlaL	177
QY	748	TGCGCGTGGCGGGAAGATTTGGGCTGTGGAAGGACATCTTCAACCGTCTCCATGAATG	807
Db	177	euProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSerMetAsnG	197
QY	808	AGAAGTTTGACTTGGTGTTATAACAGAAAGTGGGAAAGTGTGACCTCAGCTTGACTTTT	867
Db	197	lulysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPheAspPheT	217
QY	868	ATTTA 872	
Db	217	yrLeu 218	
RESULT 9			
ABR01861			
ID	ABR01861 standard; protein; 171 AA.		
XX	AC ABR01861;		
XX	19-MAY-2003 (first entry)		
DT	Human cancer-related protein, 192P2G7 v.3.		
DE	Human; cytostatic; vaccine; cancer; immune response.		
XX			
XX			
XX			
OS	Homo sapiens.		
XX	WO200283921-A2.		
XX	24-OCT-2002.		
PD			
XX	10-APR-2002; 2002WO-US011654.		
PF			
XX	10-APR-2001; 2001US-0282739P.		
PR	10-APR-2001; 2001US-0283112P.		
PR	25-APR-2001; 2001US-0286630P.		
XX	(AGEN-) AGENSYS INC.		
PA			
XX	Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;		
PI	Morrison K, Morrison RK, Raitano AB;		
PI			
XX	WPI; 2003-075555/07.		
DR	N-PSDB; ABZ78168.		
DR			
XX	New composition comprising a substance that modulates the structure of		
PT	proteins and polynucleotides, useful for therapeutic, prognostic and		
PT	diagnostic reagents for eliciting cellular or humoral immune response in		
PT	cancer patients.		
XX			
XX	Example 5; Fig 12r; 1021pp; English.		
PS			
XX	The present invention relates to novel human cancer-related genes and		
CC	proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and		
CC	proteins are useful for eliciting a humoral or cellular immune response.		
CC	The genes are useful as probes and primers for the amplification and/or		
CC	detection of genes, mRNAs or their fragments, as reagents for the		
CC	diagnosis and/or prognosis of cancer, as coding sequences capable of		
CC	directing the expression of the protein, as tools for modulating or		
CC	inhibiting the expression of genes and/or translation of transcripts, and		
CC	as therapeutic agents. The proteins and peptides are useful as		
CC	therapeutic, prognostic and diagnostic reagents for cancer		
XX			
SQ	Sequence 171 AA;		
Alignment Scores:			
Pred. No.:	3,48e-73	Length:	171
Score:	867.50	Matches:	170
Percent Similarity:	60.2%	Conservative:	1
Best Local Similarity:	59.9%	Mismatches:	0
Query Match:	19.7%	Indels:	113
DB:	6	Gaps:	1
US-10-768-158-1 (1-2419) x ABR01861 (1-171)			
QY	21	ATGCGGAGAGCGAGCGCGAGACCCCGAGACCCCGGGGAGTTTCGAGAGCAAGTACTTC	80
Db	1	MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerLysTyrPhe	20
QY	81	GAGTTCATGCGCTGCGGCTGCGCCCTTCTGCGCGGGAAGATGGAGGAGATCGCCAAC	140
Db	21	GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluGluIleAlaAsn	40
QY	141	TTCCCGGTGCGGCCAGCGAGCGTGTGGATCGTCACTACCCCAAGTCCGCCACCGCTTG	200
Db	41	PheProValArgProSerAspValTrpIleValThrTrpProLysSer-----	56
QY	201	CTGCAGGAGGTGCTCTACTTGTGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATG	260
Db	56	-----	56
QY	261	AACATCGAGCAGCAGCTCCCGTCTCGGAGTACCCACAGCCGGCGCTGGACATCATCAAG	320
Db	56	-----	56
QY	321	GAAGTACCTCTCCCGCTCATCAAGAGCCACCTGCCCTTACCGCTTCTGCGCTCTGAC	380
Db	56	-----	56

QY 381 CTCCACATGGAGCTCCAGGTCTATATGCTCGCAACCCCAAGGATCTGGTGGT 440  
 Db 56 ----- 56  
 QY 441 TCTTATTATCAGTTCACCGCTCTCTCGGACCATGAGCTACCGAGGCACCTTCAAGAA 500  
 Db 56 ----- 56  
 QY 501 TTCTGCCGAGGTTTATGAATGATAAGCTGGGTACCGCTCCGTGGTTGAGCAGTCAG 560  
 Db 57 -----ValGlyTyrGlySerTrpPheGluHisValGln 67  
 QY 561 GAGTCTCTGGAGCACCGCATGGACTCGAAGCTGCTTTTCTCAAGTATGAGCATGCAT 620  
 Db 68 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHis 87  
 QY 621 CGGACCTCGTGCAGTGTGGAGCAGCTGCCAGATTCCTGGGGGTGCTCTGTGACAAG 680  
 Db 88 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 107  
 QY 681 GCCAGCTGGAAGCCCTGACGGAGCATGTCACAGCTGTGTGACACAGTGTGCAACGCT 740  
 Db 108 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysAsnAla 127  
 QY 741 GAGCCCTGCGCTGGCGCGGAGAGTGGCTGTGGAGGCATCTTCACCGTCTCC 800  
 Db 128 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspPheThrValSer 147  
 QY 801 ATGAATCAGAAGTTTGACTTGGTGTATAAACAGAGATGGAAAGTGTGACCTCACGTTT 860  
 Db 148 MetAsnGlnLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 167  
 QY 861 GACTTTTATTTA 872  
 Db 168 AspPheTyrLeu 171  
 RESULT 10  
 ID ADZ21934  
 XX AC ADZ21934 standard; protein; 171 AA.  
 XX DT 16-JUN-2005 (first entry)  
 XX DE Human SULT4A1 splice variant 2 SEQ ID NO 2.  
 XX KW neuroprotective; nootropic; diagnosis; prognosis; gene expression;  
 KW neurodegenerative disease; neurological disease; Alzheimers disease;  
 KW degeneration; sulfotransferase family 4A member 1; SULT4A1;  
 KW splice variant; enzyme.  
 XX OS Homo sapiens.  
 XX PN WO2005030947-A2.  
 XX PD 07-APR-2005.  
 XX PF 29-SEP-2004; 2004WO-EP052353.  
 XX PR 30-SEP-2003; 2003US-0506775P.  
 XX PA (EVOT-) EVOTEC NEUROSCIENCES GMBH.  
 XX PI Von Der Kammer H, Pohlner J;  
 XX DR WPI; 2005-262951/27.  
 XX DR N-PSDB; ADZ21936.  
 XX PT Diagnosing or prognosticating a neurodegenerative disease in a subject by  
 PT determining a level and/or activity of a transcription and/or translation  
 PT product of a gene coding for a cytosolic sulfotransferase family 4A  
 member 1.

XX

Claim 15; SEQ ID NO 2; 83pp; English.

XX

The invention describes a method of diagnosing or prognosticating a neurodegenerative disease in a subject, or determining whether a subject is at increased risk of developing the disease. The method comprises determining a level and/or an activity of a transcription product and/or a translation product of a gene coding for a cytosolic sulfotransferase family 4A member 1 and/or a fragment, or derivative or variant of the transcription or translation product, in a sample obtained from the subject and comparing the level and/or the activity to a reference value representing a known disease or health status, therefore diagnosing or prognosticating the neurodegenerative disease in the subject, or determining whether the subject is at increased risk of developing the neurodegenerative disease. Protein molecules comprising fully defined 284 and/or 171 amino acid (SEQ ID NO. 1 and/or 2, respectively) sequence given in the specification, the protein molecules being translation products of the gene coding for a cytosolic sulfotransferase family 4A member 1, or its fragments, or derivatives, or variants, are useful as diagnostic targets for detecting a neurodegenerative disease, preferably Alzheimer's disease or as screening targets for reagents or compounds preventing, or treating, or ameliorating the disease. An antibody specifically immunoreactive with an immunogen, where the immunogen is a translation product of a gene coding for a cytosolic sulfotransferase family 4A member 1, SEQ ID NO. 1 or 2, or its fragment, or derivative, or variant is useful for detecting the pathological state of a cell in a sample obtained from a subject, comprising immunocytochemical staining of the cell with the antibody, where an altered degree of staining, or an altered staining pattern in the cell compared to a cell representing a known health status indicates a pathological state of the cell which relates to a neurodegenerative disease, preferably to Alzheimer's disease. The genetically altered non-human animal is useful for screening, testing, and validating compounds, agents, and modulators in the development of diagnostics and therapeutics to treat neurodegenerative diseases, in particular Alzheimer's disease. An agent(s) which directly or indirectly affect an activity and/or a level of a gene coding for a cytosolic sulfotransferase family 4A member 1, and/or a transcription product or a translation product of a gene coding for a cytosolic sulfotransferase family 4A member 1, and/or their fragment, or derivative, or variant is useful for treating or preventing a neurodegenerative disease, in particular Alzheimer's disease. This is the amino acid sequence of human sulfotransferase family 4A member 1 splice variant 2 (SULT4A1sv1).

SQ Sequence 171 AA;

Alignment Scores:

Pred. No.:	3,48e-73	Length:	171
Score:	867.50	Matches:	170
Percent Similarity:	60.2%	Conservative:	1
Best Local Similarity:	59.9%	Mismatches:	0
Query Match:	19.7%	Indels:	113
DB:	9	Gaps:	1

US-10-768-158-1 (1-2419) x ADZ21934 (1-171)

QY 21 ATGCGGAGAGCGAGCGCGAGACCCCGAGACCCCGGGGAGTTCGAGAGCAAGTACTTC 80

Db	1	MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerLysTyrPhe	20
----	---	--	----

QY 81 GAGTTCATGGCTGGCGCTGCCGCCCTTCTGCCGGGGAAGATGGAGGAGATGCCCAAC 140

Db	21	GluPheHisGlyValArgLeuProProPheCysArgGlyLysMetGluGluIleAlaAsn	40
----	----	--	----

QY 141 TTCCCGGTGCGCCCGAGCGAGCTGTGGATCGTCACTTACCCCAAGTCCCGCAGCAGCTTG 200

Db	41	PheProValArgProSerAspValTyrIleValThrTyrProLysSer	56
----	----	--	----

QY 201 CTGCAGGAGTGTCTACTTGTGTGAGCCGGCGCTGACCCCGATGAGATCGGCTTGATG 260

Db	56	-----	56
----	----	-------	----

QY 261 AACATGACGACGACGAGCTCCCGGTCTGTGAGTACCCACAGCCGGCGCTGGACATCATCAAG 320

```
Db 56 ----- 56
Qy 321 GAACTGACCTCTCCCGCCTCATCAAGAGCACCTGCGCTTCTGCGCCTCTGAC 380
Db 56 ----- 56
Qy 381 CTCACAATGGAGACTCCAAGGTCATCTATATGGCTCGCAACCCCAAGGATCTGGTGGTG 440
Db 56 ----- 56
Qy 441 TCTTATTATCAGTTCCACCGCTCTCTCGGGACCATGAGCTACCGAGGCACCTTTCAGAA 500
Db 56 ----- 56
Qy 501 TTCTGCGGAGGTTTATGAATGATAAGCTGGGTACCGCTCCTGGTTTGAGCAGCTGCAG 560
Db 57 -----ValGlyTyrGlySerTrpPheGluHisValGln 67
Qy 561 GAGTTCTGGGAGCACCGCATGGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATGCAT 620
Db 68 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHis 87
Qy 621 CGGAGCTGTGACGATGTTGGAGCAGCTGCGCAGATTCCTGGGGTGTCTCTGTGACAAG 680
Db 88 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 107
Qy 681 GCCCAGCTGGAAGCCCTGACGGAGCAGCTGCACACAGCTGGTGACCAAGTCTGCAACGCT 740
Db 108 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 127
Qy 741 GAGGCCCTGCCCTGGCGCCGGGAAGAGTTGGGCTGTGGAAGGACAPCTTTCACCGTCTCC 800
Db 128 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 147
Qy 801 ATCAATCAGAGCTTGACTTGGTGTATATAACAGAGAGATGGAAAGTGTGACCTCACGTTT 860
Db 148 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 167
Qy 861 GACTTTTATTATA 872
Db 168 AspPheTyrLeu 171

RESULT 11
AAU78269
ID AAU78269 standard; protein; 159 AA.
AC AAU78269;
XX
XX 05-JUN-2002 (first entry)
XX Rat nervous system-specific sulfotransferase, SULTN, protein sequence.
XX
XX Rat; nervous system-specific sulfotransferase; SULTN; sulfonation;
XX neurological disorder; multiple sclerosis; Huntington's chorea;
XX nervous system-mediated muscular atrophy; amyotrophic lateral sclerosis;
XX myasthenia gravis; Alzheimer's disease; senile dementia; Schizophrenia;
XX bipolar disorder; cytosolic sulfotransferase gene family;
XX cerebellar ataxia.
XX
XX Rattus sp.
XX
XX WO200218541-A2.
XX
XX 07-MAR-2002.
XX
XX 04-SEP-2001; 2001WO-US041986.
XX
XX 01-SEP-2000; 2000US-0229929P.
XX
XX (UYBO-) UNIV BOSTON.
XX
XX Farb DH, Martin S;
PI
```

```
XX
DR WPI; 2002-281063/32.
DR P-PSDB; AAU78269.
XX
XX New nervous system-specific sulfotransferase (SULTn) polypeptide for
PT treating neurological disorders including multiple sclerosis, nervous
PT system-mediated muscular atrophy and Huntington's chorea.
XX
XX Claim 2; Page; 46pp; English.
XX
XX The present invention relates to a new nervous system-specific
CC sulfotransferase (SULTn) polypeptide. The SULTn polypeptide of the
CC invention comprises a nervous system-specific sulfotransferase,
CC comprising a sequence of 284 amino acids, given in the specification. The
CC invention is useful for identifying an effector of nervous system-
CC specific sulfonation by contacting the polypeptide with a candidate
CC effector compound and assaying the ability of the candidate effector
CC compound to effect the sulfotransferase function of the polypeptide. The
CC invention is also useful for inhibiting nervous system-specific
CC sulfonation by contacting the polypeptide under physiological conditions
CC with a compound capable of interacting with and inhibiting the
CC sulfonation activity of the polypeptide, where the compound is an
CC antibody, peptide, polypeptide, nucleic acid, organic molecule or an
CC inorganic molecule. The polypeptide of the invention is also useful for
CC treating a neurological disorder, including multiple sclerosis, nervous
CC system-mediated muscular atrophy, amyotrophic lateral sclerosis,
CC Huntington's chorea, myasthenia gravis, Alzheimer's disease, senile
CC dementia, Schizophrenia, bipolar disorder and cerebellar ataxia. The
CC present amino acid sequence represents the rat SULTn protein of the
CC invention. SULTn is a member of the cytosolic sulfotransferase gene
CC family. Note: The sequence data for this sequence did not form part of
CC the printed specification but was obtained in electronic format from
CC GENBANK AF176343
XX
XX SQ Sequence 159 AA;
XX
XX Alignment Scores:
XX Pred. No.: 1.4e-71 Length: 159
XX Score: 850.50 Matches: 157
XX Percent Similarity: 99.4% Conservative: 2
XX Best Local Similarity: 98.1% Mismatches: 0
XX Query Match: 19.3% Indels: 1
XX DB: 5 Gaps: 1
XX
XX US-10-768-158-1 (1-2419) x AAU78269 (1-159)
XX
XX Qy 393 GACTCCAAAGTGCATCTATATGGCTCGCAACCCCAAGGATCTGGTGTCTTATTATCAG 452
XX Db 1 AspSerLysValIleTyrMetAlaArgAsnProLysAspLeuValSerTyrTrpGln 20
XX
XX Qy 453 TTCACCGCTCTCTCGCGACCATGAGCTACCGAGGCACCTTTCAAGAATTTCTGCCGAGG 512
XX Db 21 PheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGluPheCysArgArg 40
XX
XX Qy 513 TTTATGAATGATAAGCTGGGCTACGGCTCTCGTGTGTTGAGCACGTGACGAGTCTGGGAG 572
XX Db 41 PheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGlnGluPheTrpGlu 60
XX
XX Qy 573 CACCGCATGGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATGCATCGGAGCCTCGTG 632
XX Db 61 HisArgMetAspAlaAsnValLeuPheLeuLysTyrGluAspMetHisArgAspLeuVal 80
XX
XX Qy 633 ACATGCTGGAGCAGCTGGCCAGATTCTCTGGGGGTGTCTCTGTGACAAAGGCCAGCTGGAA 692
XX Db 81 ThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLysAlaGlnLeuGlu 100
XX
XX Qy 693 GCCCTGACGGAGCAGCTGCCACACAGCTGGTGACCATGTGTGCAACCGCTGAGGCCCTGCC 752
XX Db 101 SerLeu---GluHisCysHisGlnLeuValAspGlnCysCysAsnAlaGluAlaLeuPro 119
XX
XX Qy 753 GTGGGCGGGGAGAGTGTGGCTGTGGAAGGACATCTTCCCGTCTCCATGATGAGAG 812
XX Db 120 ValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSerMetAsnGluLys 139
```

Qy	813	TTTGACTTGGTGTTATAACAGAGATGGAAAGTGTGACCTCAGTTTGACTTTTATTTA	872
Db	140	PheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPheAspPheTyrLeu	159
RESULT 12			
AAU18674			
ID	AAU18674	standard; protein; 105 AA.	
XX	AC		
XX	AAU18674;		
XX			
DT	21-NOV-2001	(first entry)	
DE	Renal and cardiovascular-associated protein, Seq ID 113.		
XX			
KW	Human; antiinflammatory; neuroprotective; immunomodulator; vulnery;		
KW	cardiovascular; cystostatic; nephrotropic; antianemic; nephritis;		
KW	immunosuppressive; kidney disorder; renal failure; hypertension;		
KW	cardiovascular disorder; myocardial infarction; blood disorder; anaemia;		
KW	blood coagulation disorder; electrolyte imbalance disorder; cancer;		
KW	hyponatremia; hyperkalaemia; neoplastic disorder; nephroma;		
KW	autoimmune disease; inflammatory disease; reproductive system disorder;		
KW	endocrine disorder; neural activity; neurological disorder;		
KW	wound healing; respiratory disorder.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200155328-A2.		
XX			
PD	02-AUG-2001.		
XX			
PF	17-JAN-2001; 2001WO-US001359.		
XX			
PR	31-JAN-2000; 2000US-0179065P.		
PR	04-FEB-2000; 2000US-0180628P.		
PR	24-FEB-2000; 2000US-0184664P.		
PR	02-MAR-2000; 2000US-0186350P.		
PR	16-MAR-2000; 2000US-0189874P.		
PR	17-MAR-2000; 2000US-0190076P.		
PR	18-APR-2000; 2000US-01981123P.		
PR	19-MAY-2000; 2000US-0205515P.		
PR	07-JUN-2000; 2000US-0209467P.		
PR	28-JUN-2000; 2000US-0214886P.		
PR	30-JUN-2000; 2000US-0215135P.		
PR	07-JUL-2000; 2000US-0216647P.		
PR	07-JUL-2000; 2000US-0216860P.		
PR	11-JUL-2000; 2000US-0217487P.		
PR	11-JUL-2000; 2000US-0217496P.		
PR	14-JUL-2000; 2000US-0218290P.		
PR	26-JUL-2000; 2000US-0220963P.		
PR	26-JUL-2000; 2000US-0220964P.		
PR	14-AUG-2000; 2000US-0224518P.		
PR	14-AUG-2000; 2000US-0224519P.		
PR	14-AUG-2000; 2000US-0225213P.		
PR	14-AUG-2000; 2000US-0225214P.		
PR	14-AUG-2000; 2000US-0225266P.		
PR	14-AUG-2000; 2000US-0225267P.		
PR	14-AUG-2000; 2000US-0225268P.		
PR	14-AUG-2000; 2000US-0225270P.		
PR	14-AUG-2000; 2000US-0225447P.		
PR	14-AUG-2000; 2000US-0225477P.		
PR	14-AUG-2000; 2000US-0225757P.		
PR	14-AUG-2000; 2000US-0225758P.		
PR	18-AUG-2000; 2000US-0226275P.		
PR	22-AUG-2000; 2000US-0226681P.		
PR	22-AUG-2000; 2000US-0226686P.		
PR	23-AUG-2000; 2000US-0227182P.		
PR	23-AUG-2000; 2000US-0227009P.		
PR	30-AUG-2000; 2000US-0228944P.		
PR	01-SEP-2000; 2000US-0229287P.		
PR	01-SEP-2000; 2000US-0229343P.		
PR	01-SEP-2000; 2000US-0229344P.		
PR	01-SEP-2000; 2000US-0229345P.		
PR			
PR	05-SEP-2000; 2000US-0229503P.		
PR	05-SEP-2000; 2000US-0229513P.		
PR	06-SEP-2000; 2000US-0230437P.		
PR	06-SEP-2000; 2000US-0230438P.		
PR	08-SEP-2000; 2000US-0231242P.		
PR	08-SEP-2000; 2000US-0231243P.		
PR	08-SEP-2000; 2000US-0231244P.		
PR	08-SEP-2000; 2000US-0231413P.		
PR	08-SEP-2000; 2000US-0231414P.		
PR	08-SEP-2000; 2000US-0232080P.		
PR	08-SEP-2000; 2000US-0232081P.		
PR	12-SEP-2000; 2000US-0231968P.		
PR	14-SEP-2000; 2000US-0232397P.		
PR	14-SEP-2000; 2000US-0232398P.		
PR	14-SEP-2000; 2000US-0232399P.		
PR	14-SEP-2000; 2000US-0232400P.		
PR	14-SEP-2000; 2000US-0232401P.		
PR	14-SEP-2000; 2000US-0233063P.		
PR	14-SEP-2000; 2000US-0233064P.		
PR	14-SEP-2000; 2000US-0233065P.		
PR	21-SEP-2000; 2000US-0234223P.		
PR	21-SEP-2000; 2000US-0234274P.		
PR	25-SEP-2000; 2000US-0234997P.		
PR	25-SEP-2000; 2000US-02349		

PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249224P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0251990P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX WPI; 2001-488787/53.  
XX N-PSDB; AAS30195.  
XX  
XX New polynucleotides and polypeptides, useful for diagnosing, treating,  
PT preventing or prognosing e.g. kidney, cardiovascular, blood, electrolyte  
PT imbalance or neoplastic disorders, autoimmune diseases, cancers.  
XX  
XX Claim 1; SEQ ID NO 113; 506pp; English.  
XX  
XX The invention relates to novel nucleic acids and polypeptides useful for  
CC diagnosing, treating, preventing and/or prognosing disorders related to  
CC these polypeptides. The polynucleotides are especially useful in the  
CC diagnosis, prognosis, prevention and/or treatment of diseases which  
CC include kidney disorders (e.g. renal failure or nephritis),  
CC cardiovascular disorders (e.g. hypertension or myocardial infarction),  
CC blood disorders (e.g. anaemia or blood coagulation disorders),  
CC electrolyte imbalance disorders (e.g. hyponatraemia or hyperkalaemia),  
CC neoplastic disorders (e.g. nephroma or renal cell cancer), autoimmune  
CC diseases, cancers, inflammatory diseases, reproductive system disorders,  
CC endocrine disorders, neural activity and neurological disorders, wound  
CC healing and respiratory disorders. AAU18644-AAU18715 represent the novel  
CC human renal and cardiovascular-associated amino acid sequences of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at: ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 105 AA;  
SQ  
Alignment Scores:  
Pred. No.: 6,46-41 Length: 105  
Score: 528.00 Matches: 101  
Percent Similarity: 95.3% Conservative: 0  
Best Local Similarity: 95.3% Mismatches: 4  
Query Match: 12.0% Indels: 1  
DB: 4 Gaps: 0  
US-10-768-158-1 (1-2419) x AAU18674 (1-105)  
QY 380 CCTCCACAAATGGAGACTCCAGGTCATCATATGGCTCGCAACCCCAAGGATCTGGTGGT 439  
DB 1 ProProGlntrpArgLeuGlnGlyHisLeuTyrGlySerGlnProGlnGlySerGlyGly 20  
QY 440 GTCTATTATTCAGTTCCACCGCTCTCTCGGACCATGAGCTACCGAGGCACCTTTCAAGA 499  
DB 21 ValueLeuSerValProProLeuSerAlaAspHisGluLeuProArg\*\*\*LeuSerArg 40

QY 500 ATTCTGCCGAGAGTTTATGATGATAGCTGGGCTACGGCTCCTGGTTTGAGCAGCTGCA 559  
DB 41 IleLeuProGlu-PheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValG1 60  
QY 560 GGAGTTCTGGGAGCACCAGCATGGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATGCA 619  
DB 60 nGluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetH1 80  
QY 620 TCGGACCTGTGTCACGATGGTGGAGCAGCTGGCCAGATTCTCTGGGGGTGTCTCTGTGCAA 679  
DB 80 sArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCys\*\*\*I1 100  
QY 680 GGCCAGCTGGAGGCC 695  
DB 100 ePheGlnLeuGluAla 105  
RESULT 13  
ABU97289  
ID ABU97289 standard; protein; 105 AA.  
XX  
AC ABU97289;  
XX  
DT 30-JUL-2003 (first entry)  
XX  
DE Human polypeptide #31.  
XX  
KW Human; kidney disorder; cardiovascular disorder; arrhythmia;  
KW glomerulonephritis; urinary tract infection; chronic nephritis; anaemia;  
KW carcinoid heart disease; endocarditis; blood disorder; thrombosis;  
KW haemoglobin abnormality; electrolyte imbalance; neoplastic disorder;  
KW cancer; respiratory disorder; acute rhinitis; sinusitis; pharyngitis;  
KW neurological disorder; Alzheimer's disease; Parkinson's disease;  
KW Huntington's disease.  
XX  
OS Homo sapiens.  
XX  
PN US2003013649-A1.  
XX  
PD 16-JAN-2003.  
XX  
PF 21-NOV-2001; 2001US-00989442.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0217496P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225367P.  
PR 14-AUG-2000; 2000US-0225368P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0232402P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PR 17-JAN-2001; 2001US-00764863.  
XX  
PA (ROSE/) ROSEN C A.  
PA (RUBE/) RUBEN S M.  
PA (BARA/) BARASH S C.  
XX  
PI Rosen CA, Ruben SM, Barash SC;  
XX WPI; 2003-416807/39.  
XX N-PSDB; ACD01450.  
XX  
PT New nucleic acids and polypeptides, useful for diagnosing, prognosing,  
PT preventing and/or treating e.g. kidney, cardiovascular, blood,  
PT electrolyte imbalance, neoplastic, respiratory, or neurological diseases  
or disorders.  
XX  
XX Claim 11; Page 275; 363pp; English.  
XX  
CC The invention relates to isolated nucleic acids encoding novel  
CC polypeptides. The nucleic acids are useful for chromosome mapping, for  
CC radiation hybrid mapping, for detection of cancer, in gene therapy, for  
CC identifying individuals from minute biological samples, as an alternative  
CC to restriction fragment length polymorphism (RFLP) analysis, in forensic  
CC biology and as hybridisation probes for differential identification of  
CC tissues or cell types present in a biological sample. Compositions  
CC comprising the polynucleotides, polypeptides and antibodies specific for  
CC the polypeptides may be used in the diagnosis, prognosis, prevention  
CC and/or treatment of kidney disorders (e.g. glomerulonephritis, urinary  
CC tract infections, chronic nephritis), cardiovascular disorders (e.g.  
CC arrhythmias, carcioid heart disease, endocarditis), blood disorders  
CC (e.g. thrombosis, anaemia, haemoglobin abnormalities), electrolyte  
CC imbalance, neoplastic disorders (e.g. cancers), respiratory disorders  
CC (e.g. acute rhinitis, sinusitis, pharyngitis) and neurological disorders  
CC (e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease).  
CC Sequences ABU97259-ABU97330 represent human polypeptides of the invention  
XX  
SQ Sequence 105 AA;

Alignment Scores:  
Pred. No.: 6.4e-41 Length: 105



Score:	528.00	Matches:	101
Percent Similarity:	95.3%	Conservative:	0
Best Local Similarity:	95.3%	Mismatches:	4
Query Match:	12.0%	Indels:	1
DB:	6	Gaps:	0
US-10-768-158-1 (1-2419) x ABU97289 (1-105)			
Qy	380	CCTCCACAATGAGACTCAAGGTCATCTATATGGCTCGCAACCCCAAGGATCTGGTGGT	439
Db	1	ProProGlnTrpArgLeuGlnGlyHisLeuTyrGlySerGlnProGlnGlySerGlyGly	20
Qy	440	GTCTATTATCAGTTCACCGCTCTCTCGGACCATGAGCTACCGAGGCACCTTTCAAGA	499
Db	21	ValLeuLeuSerValProProLeuSerAlaAspHisGluLeuProArg**LeuSerArg	40
Qy	500	ATTCTGCGGAGGTTTATGAATGATAGCTACGGCTACGGCTCCCTGGTTTGGAGCAGCTGCA	559
Db	41	IleLeuProGlu-PheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGI	60
Qy	560	GGAGTTCGGGAGCACCGCATGGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATGCA	619
Db	60	nGluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHi	80
Qy	620	TCGGACCTGTGAGCATGGTGGAGCAGCTGGCCAGATTCTCTGGGGTGCTCTGTGCAA	679
Db	80	sArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCys***I	100
Qy	680	GGCCACGCTGGAAGCC	695
Db	100	ePheGlnLeuGluAla	105
RESULT 14			
ID	AD117131	standard; protein; 296 AA.	
XX	AD117131		
AC	AD117131;		
XX	XX		
DT	15-APR-2004	(first entry)	
XX	XX		
DE	Murine NOVX protein homologue SeqID 667.		
XX	mouse; murine; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;		
KW	inflammation; autoimmune disorder; allergy; blood disorder;		
KW	acquired immunodeficiency syndrome; AIDS; obesity; asthma;		
KW	immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;		
KW	Alzheimer's disease; infect.		
XX	Mus musculus.		
OS	WO200268649-A2.		
XX	XX		
PPN	06-SEP-2002.		
XX	XX		
PD	31-JAN-2002; 2002WO-US002785.		
XX	XX		
PF	31-JAN-2001; 2001US-0265395P.		
XX	31-JAN-2001; 2001US-0265412P.		
PR	31-JAN-2001; 2001US-0265514P.		
PR	31-JAN-2001; 2001US-0265517P.		
PR	02-FEB-2001; 2001US-0266406P.		
PR	05-FEB-2001; 2001US-0266767P.		
PR	07-FEB-2001; 2001US-0266975P.		
PR	07-FEB-2001; 2001US-0267057P.		
PR	08-FEB-2001; 2001US-0267459P.		
PR	09-FEB-2001; 2001US-0267823P.		
PR	15-FEB-2001; 2001US-0268974P.		
PR	26-FEB-2001; 2001US-0271664P.		
PR	27-FEB-2001; 2001US-0271839P.		
PR	27-FEB-2001; 2001US-0271855P.		
PR	02-MAR-2001; 2001US-0272788P.		
PR	02-MAR-2001; 2001US-0273046P.		
PR	14-MAR-2001; 2001US-0275925P.		
PR	14-MAR-2001; 2001US-0275947P.		
PR	14-MAR-2001; 2001US-0275950P.		
PR	14-MAR-2001; 2001US-0275989P.		
PR	15-MAR-2001; 2001US-0276448P.		
PR	15-MAR-2001; 2001US-0276450P.		
PR	16-MAR-2001; 2001US-0276397P.		
PR	16-MAR-2001; 2001US-0276768P.		
PR	20-MAR-2001; 2001US-0278652P.		
PR	26-MAR-2001; 2001US-0278775P.		
PR	29-MAR-2001; 2001US-0279882P.		
PR	29-MAR-2001; 2001US-0279884P.		
PR	30-MAR-2001; 2001US-0280147P.		
PR	11-APR-2001; 2001US-0282992P.		
PR	11-APR-2001; 2001US-0283083P.		
PR	20-APR-2001; 2001US-0285133P.		
PR	23-APR-2001; 2001US-0285749P.		
PR	03-MAY-2001; 2001US-0288327P.		
PR	03-MAY-2001; 2001US-0288504P.		
PR	29-MAY-2001; 2001US-0294047P.		
PR	30-MAY-2001; 2001US-0294473P.		
PR	08-JUN-2001; 2001US-0296964P.		
PR	18-JUN-2001; 2001US-0298959P.		
PR	19-JUN-2001; 2001US-0299324P.		
PR	13-AUG-2001; 2001US-0312020P.		
PR	16-AUG-2001; 2001US-0312889P.		
PR	16-AUG-2001; 2001US-0312908P.		
PR	21-AUG-2001; 2001US-0313390P.		
PR	28-AUG-2001; 2001US-0315470P.		
PR	31-AUG-2001; 2001US-0316447P.		
PR	07-SEP-2001; 2001US-0318115P.		
PR	12-SEP-2001; 2001US-0318118P.		
PR	19-SEP-2001; 2001US-0318740P.		
PR	18-OCT-2001; 2001US-0323379P.		
PR	18-OCT-2001; 2001US-0330245P.		
PR	18-OCT-2001; 2001US-0330308P.		
PR	14-NOV-2001; 2001US-0332701P.		
XX	(CURA-) CURAGEN CORP.		
XX	Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shinkets RA;		
PI	Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;		
PI	Gierlach VL, taupier RJ, Gusev VI, Colman SD, Wolenc AR, Pena CEA;		
PI	Furtak K, Grosse WM, Alsbrook JP, Lepley DM, Rieger DK, Burgess CE;		
XX	WPI; 2002-706998/76.		
DR	New NOVX polypeptides and nucleic acids, useful for preventing or		
XX	treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,		
PT	atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or		
PT	pharmacogenomics.		
XX	Disclosure; SEQ ID NO 667; 1498pp; English.		
PS	This invention relates to a novel nucleic acids, and encoded polypeptides		
XX	thereof, which have properties related to the stimulation of biochemical		
CC	or physiological responses in a cell, tissue, organ or organism.		
CC	Specifically, it refers to the use of biologically active fragments for		
CC	diagnostic and prognostic assays and furthermore in the treatment of		
CC	diverse pathological conditions. The present invention describes novel		
CC	human and murine NOVX proteins, as well as methods to modulate their		
CC	expression using antisense oligos, ribozymes and peptide nucleic acids.		
CC	The NOVX polypeptides, polynucleotides and antibodies are useful in		
CC	treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,		
CC	atherosclerosis, cancer and diabetes. Furthermore, they may be used in		
CC	treating or preventing diseases such as inflammation, autoimmune		
CC	disorders, allergies, blood disorders, acquired immunodeficiency syndrome		
CC	(AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,		
CC	arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy		
CC	and epilepsy. Accordingly, these molecules have many activities including		
CC	cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,		
CC	haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,		
CC	antiasthmatic, nephrotropic, antiarthritic, hepatotropic,		



PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249218P.
PR	06-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249264P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249300P.
PR	08-SEP-2000;	2000US-0232081P.	PR	01-DEC-2000;	2000US-0250391P.
PR	12-SEP-2000;	2000US-0231988P.	PR	01-DEC-2000;	2000US-0251160P.
PR	14-SEP-2000;	2000US-0232397P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232399P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0232400P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251869P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251989P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251990P.
PR	14-SEP-2000;	2000US-0233065P.	PR	11-DEC-2000;	2000US-0254097P.
PR	21-SEP-2000;	2000US-0234223P.	PR	05-JAN-2001;	2001US-0259678P.
PR	25-SEP-2000;	2000US-0234997P.	XX		
PR	25-SEP-2000;	2000US-0234998P.	XX		
PR	26-SEP-2000;	2000US-0235484P.	PA	(HUMA-) HUMAN GENOME SCI INC.	
PR	27-SEP-2000;	2000US-0235834P.	XX		
PR	27-SEP-2000;	2000US-0235836P.	PI	Rosen CA, Barash SC, Ruben SM;	
PR	29-SEP-2000;	2000US-0236327P.	XX		
PR	29-SEP-2000;	2000US-0236367P.	DR	WPI; 2001-541565/60.	
PR	29-SEP-2000;	2000US-0236368P.	DR	N-PSDB; ABA12558.	
PR	29-SEP-2000;	2000US-0236369P.	XX		
PR	29-SEP-2000;	2000US-0236802P.	PT	Nucleic acids encoding 3224 human nervous system antigen polypeptides,	
PR	02-OCT-2000;	2000US-0237037P.	PT	useful for preventing, diagnosing and/or treating nervous system cancers	
PR	02-OCT-2000;	2000US-0237038P.	XX	and metastases.	
PR	02-OCT-2000;	2000US-0237039P.	XX		
PR	02-OCT-2000;	2000US-0237040P.	PS	Claim 11; SEQ ID NO 4889; 1701pp + Sequence Listing; English.	
PR	13-OCT-2000;	2000US-0239935P.	XX		
PR	20-OCT-2000;	2000US-0240960P.	CC	The invention relates to novel genes (ABA11004-ABA21534) and proteins	
PR	20-OCT-2000;	2000US-0241785P.	CC	(ABA14678-ABB18001) useful for preventing, treating or ameliorating	
PR	20-OCT-2000;	2000US-0241786P.	CC	medical conditions e.g. by protein or gene therapy. The genes are	
PR	20-OCT-2000;	2000US-0241787P.	CC	isolated from a range of human tissues disclosed in the specification.	
PR	20-OCT-2000;	2000US-0241808P.	CC	The nucleic acids, proteins, antibodies and (ant)agonists are useful in	
PR	20-OCT-2000;	2000US-0241809P.	CC	the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and	
PR	20-OCT-2000;	2000US-0241826P.	CC	ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,	
PR	20-OCT-2000;	2000US-0241826P.	CC	breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune	
PR	20-OCT-2000;	2000US-0242221P.	CC	disorders e.g. Addison's disease, allergies, autoimmune haemolytic	
PR	01-NOV-2000;	2000US-0244617P.	CC	anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,	
PR	08-NOV-2000;	2000US-0246474P.	CC	multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)	
PR	08-NOV-2000;	2000US-0246475P.	CC	cardiovascular disorders such as myocardial ischaemias; (d) wound healing	
PR	08-NOV-2000;	2000US-0246476P.	CC	; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)	
PR	08-NOV-2000;	2000US-0246477P.	CC	infectious diseases such as viral, bacterial, fungal and parasitic	
PR	08-NOV-2000;	2000US-0246478P.	CC	infections. Note: The sequence data for this patent did not form part of	
PR	08-NOV-2000;	2000US-0246523P.	CC	the printed specification, but was obtained in electronic format directly	
PR	08-NOV-2000;	2000US-0246524P.	CC	from WIPO at ftp.wipo.int/pub/published_pt_sequences	
PR	08-NOV-2000;	2000US-0246525P.	XX		
PR	08-NOV-2000;	2000US-0246526P.	SQ	Sequence 102 AA;	
PR	08-NOV-2000;	2000US-0246527P.			
PR	08-NOV-2000;	2000US-0246528P.			
PR	08-NOV-2000;	2000US-0246532P.			
PR	08-NOV-2000;	2000US-0246533P.			
PR	08-NOV-2000;	2000US-0246609P.			
PR	08-NOV-2000;	2000US-0246610P.			
PR	08-NOV-2000;	2000US-0246611P.			
PR	08-NOV-2000;	2000US-0246613P.			
PR	17-NOV-2000;	2000US-0249207P.			
PR	17-NOV-2000;	2000US-0249208P.			
PR	17-NOV-2000;	2000US-0249210P.			
PR	17-NOV-2000;	2000US-0249211P.			
PR	17-NOV-2000;	2000US-0249212P.			
PR	17-NOV-2000;	2000US-0249213P.			
PR	17-NOV-2000;	2000US-0249214P.			
PR	17-NOV-2000;	2000US-0249215P.			
<hr/>					
PR	17-NOV-2000;	2000US-0249216P.	PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.	PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.	PR	17-NOV-2000;	2000US-0249264P.
PR	17-NOV-2000;	2000US-0249265P.	PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.	PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250391P.	PR	01-DEC-2000;	2000US-0250391P.
PR	01-DEC-2000;	2000US-0251160P.	PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.	PR	05-DEC-2000;	2000US-0251988P.
PR	06-DEC-2000;	2000US-0251479P.	PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.	PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251869P.	PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.	PR	08-DEC-2000;	2000US-0251989P.
PR	11-DEC-2000;	2000US-0254097P.	PR	11-DEC-2000;	2000US-0254097P.
XX			XX		
PA	(HUMA-) HUMAN GENOME SCI INC.		XX		
XX			XX		
PI	Rosen CA, Barash SC, Ruben SM;		XX		
XX			XX		
DR	WPI; 2001-541565/60.		XX		
DR	N-PSDB; ABA12558.		XX		
XX			XX		
PT	Nucleic acids encoding 3224 human nervous system antigen polypeptides,		XX		
PT	useful for preventing, diagnosing and/or treating nervous system cancers		XX		
XX			XX		
PS	Claim 11; SEQ ID NO 4889; 1701pp + Sequence Listing; English.		XX		
XX			XX		
CC	The invention relates to novel genes (ABA11004-ABA21534) and proteins		CC	(ABA14678-ABB18001) useful for preventing, treating or ameliorating	
CC	medical conditions e.g. by protein or gene therapy. The genes are		CC	isolated from a range of human tissues disclosed in the specification.	
CC	The nucleic acids, proteins, antibodies and (ant)agonists are useful in		CC	the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and	
CC	ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,		CC	breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune	
CC	disorders e.g. Addison's disease, allergies, autoimmune haemolytic		CC	anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,	
CC	multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)		CC	cardiovascular disorders such as myocardial ischaemias; (d) wound healing	
CC	; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)		CC	infectious diseases such as viral, bacterial, fungal and parasitic	
CC	infections. Note: The sequence data for this patent did not form part of		CC	the printed specification, but was obtained in electronic format directly	
CC	from WIPO at ftp.wipo.int/pub/published_pt_sequences		CC		
XX			XX		
SQ	Sequence 102 AA;		SQ	Sequence 102 AA;	
<hr/>					
Alignment Scores:			Alignment Scores:		
Pred. No.:	4,61e-36	Length:	102		
Score:	477.00	Matches:	92		
Percent Similarity:	91.1%	Conservative:	0		
Best Local Similarity:	91.1%	Mismatches:	8		
Query Match:	10.8%	Indels:	1		
DB:	4	Gaps:	0		
<hr/>					
US-10-768-158-1 (1-2419) x ABB16232 (1-102)					
Qy	1823	CCGTGGGAGGACCATTTTGGAGGTAGCCCTGTTTCACTCGGATCAGGTGGACGGCC	1882		
Db	3	ProLeuGluAspHisPheAlaSerValAlaLeuPheHisSerAspGlnVal-AlaArgPr	22		
Qy	1883	GCCTGGGTGTCTGTCACCTCATCCCTCCCTGTTATCTGAGGGAGTAAGGTGAGTCTTT	1942		

Db 22 oProAlaCysLeuSerThrSerSerLeuArgValSerGluGlyValLysValArgSerLe 42

Qy 1943 ATTGCTTCACTGCTTAATTTCTACCCACATTCGCTGAACGATGAGAGTCGGGGCC 2002

Db 42 uLeuLeuHisCysLeuilePheSerProThrIleProGluAlaMetGluSerArgGlyPr 62

Qy 2003 AGTAGCCAGCCAAACCCCGTGGGACCGGGTGTCTGTCATTATGTGGCTGGAAAGCAC 2062

Db 62 oValAlaSerProThr\*\*\*GlyAspArgGlyCysLeu\*\*\*PheMetIrpLeuGluSerTh 82

Qy 2063 CCAAAGTGGTGGTCAGAGGGTCTGCTGTGGAAGGGTCTCCGTTCTTGGTCTGTAT 2122

Db 82 rGlnSerGlyGlyGlnGlu\*\*\*ArgCysCysGlyArgGlyLeuArgSerIrpCysCysIl 102

Qy 2123 T 2123

Db 102 e 102

RESULT 16

ID ABG27554 standard; protein; 1305 AA.

XX AC ABG27554;

DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #27545.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KM food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX WO200175067-A2.

XX FN 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HVSE-) HVSEQ INC.

XX PI Dmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS91741.

XX PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX Claim 20; SEQ ID NO 57913; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 1305 AA;

Alignment Scores: 3.44e-35 Length: 1305

Pred. No.: 474.00 Matches: 110

Score: 50.0% Conservative: 54

Best Local Similarity: 33.5% Mismatches: 110

Query Match: 10.8% Indels: 54

DB: 4 Gaps: 8

US-10-768-158-1 (1-2419) x ABG27554 (1-1305)

Qy 27 GAGAGCGAGCGGAGAGACCCCGACCCCG----- 56

Db 265 GluGluGlyArgGluGlnProTyrThrProTyrThrProIleLeuSerProCysArg 284

Qy 57 ---GGGAGTTCGAGACAAAGTACTTCGAGTTCATGGCGTGGCGGTG----- 101

Db 285 GlnGlyGluAsnArgSerGlySerArgGluLeuArgAsnMetGluLeuIleGlnAspThr 304

Qy 102 -----CCGCC-----TTCGCGCGGG-----AAG 122

Db 305 SerArgProLeuGluTyrValLysGlyValProLeuIleLysTyrPheAlaGluAla 324

Qy 123 ATGAGGAGATCGCAACTTCCCGGTGCGGCCGAGCAGCGTGGATCGTCACTACCCC 182

Db 325 LeuGlyProLeuGlnSerPheGlnAlaArgProAspLeuLeuIleAsnThrTyrPro 344

Qy 183 AAGTCGGCACCACTTGTGCGAGAGTGTCTACTTGTGTGAGCCAGGCGGTGACCCC 242

Db 345 LysSerGlyThrThrTrpValSerGlnIleLeuAspMetIleTyrGlnGlyGlyAspLeu 364

Qy 243 GATGAGATCGGCTTGATGAACATCGAGCAGCAGCTCCCGTCTGGAGTACCCACGCCG 302

Db 365 GluLysCysAsnArgAlaProIleTyrValArgValProPheLeuGluValAsnAspPro 384

Qy 303 -----GGCTGGACATCATCAAGAACTGACCTCTCCCGCCCTCATCAAGAGC 350

Db 385 GlyGluProSerGlyLeuGluThrLeuLysAspThrProProArgLeuIleLysSer 404

Qy 351 CACTGCCCTACCGCTTTCGCTCTGACCTCACAATGGAGACTCAAGGTCACTAT 410

Db 405 HisLeuProLeuAlaLeuLeuProGlnThrLeuLeuAspGlnLysValLysValValTyr 424

Qy 411 ATGGCTCGCAACCCAGGATCTGGTGTCTCTATTATCAGTTCACCGCTCTCTCGGG 470

Db 425 ValAlaArgAsnProLysAspValAlaValSerTyrTyrHisPheHisArgMetGluLys 444

Qy 471 ACCATGAGCTACCGAGCACCTTTCAAGAAATTTCGCGGAGGTTTATGAATGAAGCTG 530

Db 445 AlaHisProGluProGlyThrTrpAspSerPheLeuGluLysPheMetAlaGlyVal 464

Qy 531 GGCTACGGCTCTGGTTTGGAGCAGCTGAGAGTTCCTGGAGACCCCGATCGAGTCAAC 590

Db 465 SerTyrGlySerTrpTyrGlnHisValGlnGluTrpGluLeuSerArgThrHisPro 484

Qy 591 GTGCTTTTTCAGTATGAAGACATCATCGGACCTGTCGATGGTGGAGCAGCTG 650

Db 485 ValLeuLeuPheTyrGluAspMetLysGluAsnProLysArgGluIleGlnLysIle 504

Qy 651 GCCAGATTCTCGGGGTGTCTCTGTGACAAGCCCGAGCTGGAAGCCCTGACGGAGCTGC 710

Db 505 LeuGluPheValGlyArgSerLeuLeuProGluGluThrMetAspPheMetValGlnHisThr 524

Qy 711 -----CACCAGCTG 719

Db 525 SerPheLysGluThrLysLysAsnProMetThrAsnTyrThrThrValProGlnGluLeu 544



DT 15-JAN-2004 (first entry)  
 XX Human disease related protein SeqID111.  
 DE human; disease state; cytostatic; antiinflammatory; ophthalmological;  
 XX antiarteriosclerotic; vulnerary; gene therapy;  
 KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;  
 KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;  
 KW glucose transportation; catecholamine synthesis; iron transport;  
 KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;  
 KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;  
 KW inflammatory condition; wound healing.  
 XX Homo sapiens.  
 OS WO2003018621-A2.  
 FN 06-MAR-2003.  
 PD 23-AUG-2002; 2002WO-GB003892.  
 XX 23-AUG-2001; 2001GB-00020558.  
 PR 05-OCT-2001; 2001GB-00024037.  
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.  
 XX Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;  
 PI WPI; 2003-290046/28.  
 DR N-PSDB; ADD18681.  
 XX New substantially purified polypeptide, useful for diagnosing or treating  
 PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion  
 PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or  
 PT wound healing.  
 XX Claim 25; SEQ ID NO 111; 424pp; English.  
 PS This invention relates to novel human genes and gene product which are  
 CC implicated in certain disease states. Compounds which modulate the  
 CC proteins of the invention may have cytostatic, antiinflammatory,  
 CC ophthalmological, antiarteriosclerotic or vulnerary activities. The  
 CC sequences of the invention may be useful for gene therapy. The invention  
 CC may be useful for diagnosing or treating a hypoxia-regulated condition,  
 CC such as tumorigenesis, angiogenesis, apoptosis, inflammation,  
 CC erythropoiesis, or the biological response to hypoxia conditions  
 CC including processes such as glycolysis, gluconeogenesis, glucose  
 CC transportation, catecholamine synthesis, iron transport or nitric oxide  
 CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion  
 CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,  
 CC inflammatory conditions or wound healing. The present sequence is that of  
 CC a disease related protein of the invention.  
 XX Sequence 295 AA;  
 SQ  
 Alignment Scores:  
 Pred. No.: 2,43e-35 Length: 295  
 Score: 472.00 Matches: 100  
 Percent Similarity: 54.5% Conservative: 52  
 Best Local Similarity: 35.8% Mismatches: 101  
 Query Match: 10.7% Indels: 26  
 DB: Gaps: 4  
 US-10-768-158-1 (1-2419) x ADD18680 (1-295)  
 Qy 90 GCGGTGCGGCTCCGCCCTTCTCGCGGGAAGATGGAGAGATCGCAACTTCCCGGTG 149  
 Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36  
 Qy 150 CGGCCACGACGGTGTGGATGCTACCTACCCCAAGTCCGGACCGAGCTGCTGGAGGAG 209  
 Db 37 ArgProAspAspLeuLeuIleAsnThrTyrProLysSerGlyThrThrTrpValSerGln 56  
 Qy 210 GTGGTCTACTGGTGAGCCAGGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269  
 Db 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGlnLysCysAsnArgAlaProIleTyr 76  
 Qy 270 GAGCAGCTCCCGGTCCTGGAGTACCCACAGCCG-----GGCCTGGACATCATC 317  
 Db 77 ValArgValProPheLeuGluValAsnAspProGlyGluProSerGlyLeuGluThrLeu 96  
 Qy 318 AAGAACTGACCTCTCCCGCTCATCAAGAGCCACTCCCTACCGCTTCTCCCTCT 377  
 Db 97 LysAspThrProProArgLeuIleLysSerHisLeuProLeuAlaLeuLeuProGln 116  
 Qy 378 GACCTCCACATGGAGACTCCCAAGTCTATATGCTCGCAACCCCAAGATCTGGTG 437  
 Db 117 ThrLeuLeuAspGlnLysValLysValValAlaArgAsnProLysAspValAla 136  
 Qy 438 GTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGCTACCGAGCACCTTTCAA 497  
 Db 137 ValSerTyrTyrHisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAsp 156  
 Qy 498 GAATTCCTCGGAGGTTTATGATGATAGTACGGCTACGGCTCTCTGGTTGACACAGTG 557  
 Db 157 SerPheLeuGluLysPheMetAlaGlyGluValSerTyrGlySerTyrGlnHisVal 176  
 Qy 558 CAGGAGTCTGGGAGCACCGCATGGACTCGAAACGTCTTTTCTCAAGTATGAAGACATG 617  
 Db 177 GlnGluTrpGluLeuSerArgThrHisProValLeuLysPheThrPheThrGluAspMet 196  
 Qy 618 CATCGGACCTGGTGACGATGGTGACGATGGCCAGATTCCTGGGGGTGCTCTGTGAC 677  
 Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216  
 Qy 678 AAGCCCCAGCTGGAGCCCTGACGGACACTGC----- 710  
 Db 217 GluGluThrMetAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236  
 Qy 711 -----CACAGCTGGTGACCAAGTCTGCTGCAACGCTGAGGCC 746  
 Db 237 MetThrAsnTyrThrThrValProGlnGluLeuMetAspHisSerIleSer----- 253  
 Qy 747 CTGCCC---GTGGCCCGGGAAGAGTTGGGTGGGAAGGACATCTTTCACCGTCTCCATG 803  
 Db 254 ---ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGln 272  
 Qy 804 AATGAGAAGTTGACTTGGTGTATAAACAGAGATGGAAAGTGTGACCTCACGTTT 860  
 Db 273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291  
 RESULT 19  
 ADP24302  
 ID ADP24302 standard; protein; 295 AA.  
 XX  
 AC ADP24302;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE PRO polypeptide SEQ ID NO:1480.  
 XX  
 KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;  
 KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;  
 KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.  
 XX Unidentified.  
 OS  
 XX WO2004041170-A2.  
 FN  
 XX 21-MAY-2004.  
 PD  
 XX 30-OCT-2003; 2003WO-US034312.  
 PF  
 XX 01-NOV-2002; 2002US-0423394P.  
 PR  
 XX (GETH ) GENENTECH INC.  
 PA



Genetically typing animals to determine those with desired boar taint characteristics, comprises obtaining a sample of genetic material from the animal, and assaying for the presence of a sulfotransferase allele.

Example; SEQ ID NO 4; 24pp; English.

The present invention relates to a method of genetically typing animals to determine those with desired boar taint characteristics. The method involves obtaining a sample of genetic material from the animal and assaying for the presence of a sulfotransferase allele. The present sequence is the human phenol sulfotransferase (PST) protein, SULT1A3. The SULT1A3 protein encoding gene is located on chromosome 16p12.1.

Sequence 295 AA;

Qy	747	CTGCCC---	GTGGCCCGGGGAAGAGTTGGGCTGTGGGAAGACATCTTCACCGTCTCCATG	808
Db	254	---PropHeMet	ArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGln	272
Qy	804	AATGAGAGTTTGAC	TTCGTTGATATAACAGAGAAGATGGAAAAGTGTGACCTCAGCTT	860
Db	273	AenGluAArgPhe	AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe	291
RESULT 21				
ADI117130				
ID	ADI117130	standard; protein; 296 AA.		
AC	ADI117130;			
XX	XX			
DT	15-APR-2004	(first entry)		
DE	Rat NOVX	protein homologue SeqID 666.		
KW	rat; NOVX;	cardiomyopathy; atherosclerosis; cancer; diabetes;		
KW	inflammation;	autoimmune disorder; allergy; blood disorder;		
KW	acquired immunodeficiency syndrome; AIDS;	obesity; asthma;		
KW	immunoglobulin (Ig)A nephropathy; cirrhosis;	arthritis;		
KW	Alzheimer's disease;	infection; str.		
OS	Rattus norvegicus.			
XX	XX			
FN	WO200268649-A2.			
PD	06-SEP-2002.			
XX	XX			
PF	31-JAN-2002;	2002WO-US002785.		
XX	XX			
PR	31-JAN-2001;	2001US-0265395P.		
PR	31-JAN-2001;	2001US-0265412P.		
PR	31-JAN-2001;	2001US-0265514P.		
PR	02-FEB-2001;	2001US-0265517P.		
PR	02-FEB-2001;	2001US-0266406P.		
PR	05-FEB-2001;	2001US-0266767P.		
PR	07-FEB-2001;	2001US-0266975P.		
PR	08-FEB-2001;	2001US-0267057P.		
PR	09-FEB-2001;	2001US-0267459P.		
PR	15-FEB-2001;	2001US-0267823P.		
PR	26-FEB-2001;	2001US-0268974P.		
PR	27-FEB-2001;	2001US-0271664P.		
PR	27-FEB-2001;	2001US-0271839P.		
PR	02-MAR-2001;	2001US-0271855P.		
PR	02-MAR-2001;	2001US-0272788P.		
PR	14-MAR-2001;	2001US-0273046P.		
PR	14-MAR-2001;	2001US-0275925P.		
PR	14-MAR-2001;	2001US-0275947P.		
PR	14-MAR-2001;	2001US-0275950P.		
PR	14-MAR-2001;	2001US-0275989P.		
PR	15-MAR-2001;	2001US-0276448P.		
PR	15-MAR-2001;	2001US-0276450P.		
PR	16-MAR-2001;	2001US-0276397P.		
PR	16-MAR-2001;	2001US-0276768P.		
PR	20-MAR-2001;	2001US-0278652P.		
PR	26-MAR-2001;	2001US-0278775P.		
PR	26-MAR-2001;	2001US-0278788P.		
PR	29-MAR-2001;	2001US-0279882P.		
PR	29-MAR-2001;	2001US-0279884P.		
PR	30-MAR-2001;	2001US-0280147P.		
PR	11-APR-2001;	2001US-0282992P.		
PR	11-APR-2001;	2001US-0283083P.		
PR	20-APR-2001;	2001US-0285133P.		
PR	23-APR-2001;	2001US-0285749P.		
PR	03-MAY-2001;	2001US-0288327P.		
PR	03-MAY-2001;	2001US-0288504P.		
PR	29-MAY-2001;	2001US-0294047P.		
PR	30-MAY-2001;	2001US-0294473P.		
PR	08-JUN-2001;	2001US-0296964P.		
PR	18-JUN-2001;	2001US-0298959P.		







CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX  
SQ Sequence 291 AA;

Alignment Scores:  
Pred. No.: 5.21e-34 Length: 291  
Score: 458.00 Matches: 96  
Percent Similarity: 55.6% Conservative: 57  
Best Local Similarity: 34.9% Mismatches: 104  
Query Match: 10.4% Indels: 18  
DB: 7 Gaps: 4

US-10-768-158-1 (1-2419) x ADD45448 (1-291)

QY 90 GGCGTGGCGTCCGCGCTTCTGCGCGGGAAGATGGAGGATCGCCACTTCCCGTG 149  
Db 13 GlylleProleuLleLysTyPheAlaGluThrIleGlyProLeuGlnAsnPheThAla 32  
QY 150 CGGCCCGCAGCGTGTGGATCGTCACCTACCCCAAGTCCGGCAGCTGCTCGCAGGAG 209  
Db 33 TrpProaspPheLeuLleSerThrTyProLysSerGlyThrTrpMetSerGlu 52  
QY 210 GTGTCTACTTGTGAGCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATGCAC 269  
Db 53 IleLeuAspMetIleTyGlnGlyGlyLeuGluLysCysGlyArgAlaProIleTy 72  
QY 270 GAGCAGCTCCCGTCCGAGTACCACAGCCG-----GGCCTGGACATCATC 317  
Db 73 AlaargValProPheLeuGluPheLysCysProGlyValProSerGlyLeuGluThrLeu 92  
QY 318 AAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTCTGCCCTCT 377  
Db 93 GluGluThrProAlaProArgLeuLeuLysThrHisLeuProLeuSerLeuLeuProGln 112  
QY 378 GACCTCCACATGAGACTCCAAAGTCAATCATATGCTCGCAACCCCAAGATCTGGTG 437  
Db 113 SerLeuLeuAspGlnLysValLysValIleTyIleAlaArgAsnAlaLysAspValVal 132  
QY 438 GTGTCTTATTATCATGTTCCACCGCTCTCTCGGACCATGAGTACCGAGGCACCTTTCAA 497  
Db 133 ValSerTy 152  
QY 498 GAATTCCTCCCGAGGTTTATGAATGATAAGCTGGGCTACGGCTCCTGGTTTGACACGTG 557  
Db 153 SerPheLeuGluAsnPheMetAspGlyGluValSerTyTyTyTyTyTyTyTyTyTyTyTy 172  
QY 558 CAGAGTTCTGGGAGCACCAGCATGAGTCAACCTGCTTTCTCAAGTATGAAGACATG 617  
Db 173 LysGluTrpTrpGluLeuArgHisThrHisProValLeuTyLeuPheTyTyGluAspIle 192  
QY 618 CATCGGACCTGTGACATGTTGGAGCAGCTGCCAGATTCCTGGGGGTGCTCTGTGAC 677  
Db 193 LysGluAsnProLysArgGluIleLysLysIleLeuGluPheLeuGlyArgSerLeuPro 212  
QY 678 AAGGCCCGAGCTGGAGCCCTGACGAGCATGCG-----CACCAAGCTGGTGGACCATGCG 731  
Db 213 GluGluThrValAspSerIleValHisThrSerPheLysLysMetLysGluAsnCys 232

QY 732 ---TGCACCGCTGAGCCCTGCCC-----GTG 755  
Db 233 MetThrAsnTyTyThrThrIleProThrGluIleMetAspHisAsnValSerProPheMet 252  
QY 756 GGCCGGGGAAGAGTTGGCTGTGGAAGGACATCTTACCGTCTCATGAATGAGAAGTTT 815  
Db 253 ArgLysGlyThrThrGlyAspTrpLysAsnThrPheThrValAlaGlnAsnGluArgPhe 272  
QY 816 GACTTGTGTATAACAGAGATGGAAAGTGTGACCTCAGCTT 860  
Db 273 AspAlaHisTyTyAlaLysThrMetThrAspCysAspPheLysPhe 287  
RESULT 24  
ADE57139  
ID ADE57139 standard; protein; 291 AA.  
XX  
AC ADE57139;  
DT  
DT 29-JAN-2004 (first entry)  
XX  
XX Rat Protein P17988, SEQ ID NO 2999.  
DE  
XX  
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX  
OS Rattus norvegicus.  
XX  
XX WO2003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
XX (GEHO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
XX Woolf C, D'urso D, Befort K, Costigan M;  
XX  
XX WPI; 2003-268312/26.  
DR GENBANK; P17988.  
XX  
PT New composition comprising two or more isolated polypeptides, useful for  
XX preparing a medicament for treating pain in an animal.  
PS Claim 1; Page; 1017pp; English.  
XX  
XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of

CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 291 AA;

Alignment Scores:  
Pred. No.: 5.21e-34 Length: 291  
Score: 458.00 Matches: 96  
Percent Similarity: 55.6% Conservative: 57  
Best Local Similarity: 34.9% Mismatches: 104  
Query Match: 10.4% Indels: 18  
DB: 7 Gaps: 4

US-10-768-158-1 (1-2419) x ADE57139 (1-291)

```
QY 90 GCGCTGCGCTGCGCGCTCTCTCGCGGGGAAGATGGAGGATCGCAACTTCCCGGTG 149
D 13 GlylleProLeuIleLysTyrPheAlaGluThrIleGlyProLeuGlnAsnPheThrAla 32
QY 150 CGGCCCGACGACGTGTGGATCGTCACTACCCCAAGTCCGCCACGAGCTTGTCTCGAGAG 209
D 33 TrpProAspAspLeuLeuIleSerThrTyrProLysSerGlyThrThrTrpMetSerGlu 52
QY 210 GTGTCTACTTGGTGACCGCGGCTGACCCCGATGAGATCGCGTGTGATGAACATCGAC 269
D 53 IleLeuAspMetIleTyrGlnGlyLysLeuGluLysCysGlyA-gAlaProIleTyr 72
QY 270 GAGCAGCTCCCGCTCGAGTACCCACAGCCG-----GGCTGGACATCATC 317
D 73 AlaArgValProPheLeuGluPheLysCysProGlyValProSerGlyLeuGluThrLeu 92
QY 318 AAGAACTGACCTCTCCCGCTCATCAAGACCCACCTGCGCTTCTGCTTCTGCTTCT 377
D 93 GluGluThrProAlaProArgLeuLeuLysThrHisLeuProLeuSerLeuProGln 112
QY 378 GACCTCCCAATGAGACCTCAAGTCTATATATGGTCTCGCAACCCCAAGATCTGGTG 437
D 113 SerLeuLeuAspGlnLysValIleTyrIleAlaArgAsnAlaLysAspValVal 132
QY 438 GTGTCTTATTATCAGTTCACCGCTCTCTCGCGGACCATGAGTACCGAGGACCTTTCAA 497
D 133 ValSerTyrTyrAsnPheTyrAsnMetAlaLysLeuHisProAspProGlyThrTrpAsp 152
QY 498 GAATTCTCGCGAGGTTTATGAATGATAGTGGCTACGGCTCTCTGTTTGGACAGCTG 557
D 153 SerPheLeuGluAsnPheMetAspGlyGluValSerTyrGlySerTyrGlnHisVal 172
QY 558 CAGGAGTCTGGAGCACCGCATGACTCGAACGTGCTTTTCTCAAGTATGAGACATG 617
D 173 LysGluThrTrpGluLeuArgHisThrHisProValLeuTyrLeuPheTyrGluAspIle 192
QY 618 CATCGGACCTGGTGACCATGTGGAGCAGCTGGCCAGATTCCTGGGGGTCTCTGTGAC 677
D 193 LysGluAsnProLysArgGluIleLysLysIleLeuGluPheLeuGlyArgSerLeuPro 212
QY 678 AAGGCCACGCTGGAAGCCCTGACGAGCAGCTGC-----CACCAGCTGGTGGACAGTGC 731
D 213 GluGluThrValAspSerIleValHisHisThrSerPheLysLysMetLysGluAsnCys 232
QY 732 ---TGCAACGCTGAGCCCTGCC-----GTG 755
D 233 MetThrAsnTyrThrIleProThrGluLeuMetAspHisAsnValSerProPheMet 252
QY 756 GCGCGGGAAGAGTTGGCTGTGGAGGACATCTTCCCGCTCTCCATCAATGAGAGTTT 815
D 253 ArgLysGlyThrThrGlyAspTrpLysAsnThrPheThrValAlaGlnAsnGluArgPhe 272
QY 816 GACTTGGTGTTATAACAGAGAGTGGAAAGTGTGACCTCAGCTTT 860
D 273 AspAlaHisTyrAlaLysThrMetThrAspCysAspPheLysPhe 287
```

RESULT 25

ADE57135

ID ADE57135 standard; protein; 291 AA.

XX AC ADE57135;

XX DT 29-JAN-2004 (first entry)

XX DE Rat Protein P17988, SEQ ID NO 2995.

XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

XX KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX PN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GEHO ) GEN HOSPITAL CORP.

XX PA (FARB ) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX PT WPI; 2003-268312/26.

XX DR GENBANK; P17988.

XX SQ New composition comprising two or more isolated polypeptides, useful for

preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat

or human polynucleotides or a polynucleotide which represents a fragment,

derivative or allelic variation of the nucleic acid sequence. Also

claimed are a vector comprising the novel polynucleotide, a host cell

comprising the vector, a method for identifying a nucleotide sequence

which is differentially regulated in an animal subjected to pain and a

kit to perform the method, an array, a method for identifying an agent

that increases or decreases the expression of the polynucleotide sequence

that is differentially expressed in neuronal tissue of a first animal

subjected to pain, a method for identifying a compound which regulates

the expression of a polynucleotide sequence which is differentially

expressed in an animal subjected to pain, a method for identifying a

compound that regulates the activity of one or more of the

polynucleotides, a method for producing a pharmaceutical composition, a

method for identifying a compound or small molecule that regulates the

activity in an animal of one or more of the polypeptides given in the

specification, a method for identifying a compound useful in treating

pain and a pharmaceutical composition comprising the one or more

polypeptides or their antibodies. The polynucleotide or the compound that

modulates its activity is useful for preparing a medicament for treating

pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene

therapy). The sequence presented is a rat protein (shown in Table 2 of

the specification) which is differentially expressed during pain. Note:

The sequence data for this patent did not form part of the printed

specification, but was obtained in electronic form directly from WIPO at

ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 291 AA;

Alignment Scores:

Pred. No.: 5.21e-34 Length: 291

Score: 458.00 Matches: 96

Percent Similarity: 55.6% Conservative: 57

Best Local Similarity: 34.9% Mismatches: 104

Query Match: 10.4% Indels: 18  
DB: Gaps: 4

US-10-768-158-1 (1-2419) x ADE57135 (1-291)

QY 90 GCGTGGCGTGGCGCCCTCTCGCGGGAGATGGAGGATGCCCACTTCCCGGTG 149  
DB 13 GlylleProLeuIleLysTyrPheAlaGluThrIleGlyProLeuGlnAsnPhenThrAla 32

QY 150 CGGCCAGCAGCTGTGGATCGTCACCTACCCCAAGTCCGGACAGCTGTGCAGGAG 209  
DB 33 TrpProAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpMetSerGlu 52

QY 210 GTGCTCTACTGTGAGCCGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269  
DB 53 IleLeuAspMetIleTyrGlnGlyGlyLysLeuGluLysCysGlyArgAlaProIleTyr 72

QY 270 GAGCAGCTCCGCTCTGGAGTACCCACGCCG-----GGCTGACATCATC 317  
DB 73 AlaArgValProPheLeuGluPheLysCysProGlyValProSerGlyLeuGluThrLeu 92

QY 318 AAGGAACGTACCTCTCCCGCTCATCAAGAGCCACTGCCCTACCGCTTTCTGCCCTCT 377  
DB 93 GluGluThrProAlaProArgLeuLeuLysThrHisLeuProLeuSerLeuLeuProGln 112

QY 378 GACCTCCAAATGGAGACTCCAGTTCATCTATATGGCTCGCAACCCCAAGGATCTGGTG 437  
DB 113 SerLeuLeuAspGlnLysValLysValIleTyrIleAlaArgAsnAlaLysAspValVal 132

QY 438 GTGCTCTATATCAGTTCACCGCTCTCTCGGACCATGAGTACCGAGGACCTTTCAA 497  
DB 133 ValSerTyrTyrAsnPheTyrAsnMetAlaLysLeuHisProAspProGlyThrTrpAsp 152

QY 498 GAATTCGCGGAGGTTTATGAATGATAGCTGGGCTACGGCTCTGGTTTTCGAGCACGTG 557  
DB 153 SerPheLeuGluAsnPheMetAspGlyGluValSerTyrGlySerTrpTyrGlnHisVal 172

QY 558 CAGGAGTTCGCGGAGCACCGCATCGACTCGAACGTGCTTTTCTCAAGTATGAAGACATG 617  
DB 173 LysGluTrpTrpGluLeuArgHisThrHisProValLeuTyrLeuPheTyrGluAspIle 192

QY 618 CATCGGACCTGTGACGATGTGGAGACGCTGGCCAGATTCCTGGGGGTGTCGTGTGAC 677  
DB 193 LysGluAsnProLysArgGluIleLysIleLeuGluPheLeuGlyArgSerLeuPro 212

QY 678 AAGGCCAGCTGGAAGCCCTGACGGAGCACTGC-----CACCAGCTGTGGACCAAGTC 731  
DB 213 GluGluThrValAspSerIleValHisThrSerPheLysLysMetLysGluAsnCys 232

QY 732 ---TGCAACGCTGAGCCCTGCCC-----GTG 755  
DB 233 MetThrAsnTyrThrThrIleProThrGluIleMetAspHisAsnValSerProPheMet 252

QY 756 GGCCGGGAAGATGGGTGTGGAGGACATCTTCACCGCTTCCATGAATGAGAAGTTT 815  
DB 253 ArgLysGlyThrThrGlyAspTrpLysAsnThrPheThrValAlaGlnAsnGluArgPhe 272

QY 816 GACTTGTGTATAAACAGAAGATGGGAAGTGTGACCTCACCGTTT 860  
DB 273 AspAlaHisTyrAlaLysThrMetThrAspCysAspPheLysPhe 287

RESULT 26  
ADE57147  
ID ADE57147 standard; protein; 291 AA.  
XX AC ADE57147;  
XX DT 29-JAN-2004 (first entry)  
XX XX  
XX DE Rat Protein P17988, SEQ ID NO 3007.  
XX DE  
XX DE Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
KW

```
Qy 210 GTGCTCTACTTGGTGAGCCAGGGCGCTGACCCCGATGAGATCGCGTTCATGAAATCATCGAC 269
Db 53 ILeuAspMetIleTyrGlnGlyGlyLeuGluLysCysGlyAlaProIleTyr 72
Qy 270 GAGCAGCTCCCGGTCCTGGAGTACCACAGCCG-----GGCTGGACATCATC 317
Db 73 AlaArgValProPheLeuGluPheLysCysProGlyValProSerGlyLeuGluThrlu 92
Qy 318 AAGAACTGACTCTCCCGCTCATCAAGAGCCACCTGCGCTACCGCTTTCTGCCCTCT 377
Db 93 GluGluThrProAlaProArgLeuLeuLysThrHisLeuProLeuSerLeuLeuProGln 112
Qy 378 GACTCCACAATGAGAGATCCAAAGTCTATATATGCTCGCAACCCCAAGATCTGGTG 437
Db 113 SerLeuLeuAspGlnLysValLysValIleTyrIleAlaArgAsnAlaLysAspValVal 132
Qy 438 GTGCTTATTATAGTTCCACCGCTCTCTCGGACCATGAGCTACCGAGGACCTTCAA 497
Db 133 ValSerTyrTyrAsnPheTyrAsnMetAlaLysLeuHisProAspProGlyThrTrpAsp 152
Qy 498 GAATTCCTCCGAGGAGTTTATGAATGATAAGCTGGGCTACGGCTCTCTGGTTGAGCAGTG 557
Db 153 SerPheLeuGluAsnPheMetAspGlyGluValSerTyrGlySerTyrTrpTyrGlnHisVal 172
Qy 558 CAGGAGTCTGGGAGCACCAGCATGAGCTCGAACGTGCTTTTCTCAAGTATGAAGATG 617
Db 173 LysGluThrTrpGluLeuArgHisThrHisProValLeuTyrLeuPheTyrGluAspIle 192
Qy 618 CATCGGACCTGGTGACGATGGTGAGCAGCTGGCAGATTCCTGGGGGTCTCTGTGAC 677
Db 193 LysGluAsnProLysArgGluIleLysLysIleLeuGluPheLeuGlyArgSerLeuPro 212
Qy 678 AAGCCCAAGCTGGAGCCCTGACGAGCATGTC-----CACCACTGGTGAGCAGTGC 731
Db 213 GluGluThrValAspSerIleValHisHisThrSerPheLysLysMetLysGluAsnCys 232
Qy 732 ---TGCAACGCTGAGGCCCTCGCC-----GTG 755
Db 233 MetThrAsnTyrThrThrIleProThrGluIleMetAspHisAsnValSerProPheMet 252
Qy 756 GCGCGGGAAGATTGGGCTGTGGAGGACATCTTCACCGCTCATCGAATGAGAAGTTT 815
Db 253 ArgLysGlyThrThrGlyAspTrpLysAsnThrPheThrValAlaGlnAsnGluArgPhe 272
Qy 816 GACTTGGTGTTATAACAGAAGATGGGAAAGTGTGACCTCACCTTT 860
Db 273 AspAlaHisTyrAlaLysThrMetThrAspCysAspPheLysPhe 287
RESULT 27
ADE57143
ID ADE57143 standard; protein; 291 AA.
XX AC ADE57143;
XX DT 29-JAN-2004 (first entry)
XX DE Rat Protein P17988, SEQ ID NO 3003.
XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX OS Rattus norvegicus.
XX FN WO2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
```

XX (GEHO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX MPI: 2003-268312/26.  
XX GENBANK; P17988.

XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page: 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. Gene  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 291 AA;

Alignment Scores:  
Pred. No.: 5.21e-34 Length: 291  
Score: 458.00 Matches: 96  
Percent Similarity: 55.6% Conservative: 57  
Best Local Similarity: 34.9% Mismatches: 104  
Query Match: 10.4% Indels: 18  
DB: 7 Gaps: 4

US-10-768-158-1 (1-2419) x ADE57143 (1-291)

Qy 90 GCGTGCGGCTCGCGCTTCTGCGCGGAAGATGGAGAGATCGCAACTTCCCGTG 149  
Db 13 GlyIleProLeuIleLysTyrPheAlaGluThrIleGlyProLeuGlnAsnPheThAla 32  
Qy 150 CGGCCCGCAGCGGTGTGATCGTCACCTACCCCAAGTCCGCGACACAGCTTGTCTCGAGAG 209  
Db 33 TrpProAspLeuLeuIleSerThrTyrProLysSerGlyThrThrTrpMetSerGlu 52  
Qy 210 GTGCTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGCGTTGATGAACATCGAC 269  
Db 53 ILeuAspMetIleTyrGlnGlyGlyLysLeuGluLysCysGlyArgAlaProIleTyr 72  
Qy 270 GAGCAGCTCCCGTCTGGAGTACCACAGCCG-----GGCTGGACATCATC 317  
Db 73 AlaArgValProPheLeuGluPheLysCysProGlyValProSerGlyLeuGluThrlu 92  
Qy 318 AAGAACTGACTCTCCCGCTCATCAAGAGCCACCTGCGCTACCGCTTTCTGCCCTCT 377  
Db 93 GluGluThrProAlaProArgLeuLeuLysThrHisLeuProLeuSerLeuLeuProGln 112



Db 237 MetThrAsnTyrThrThrValArgArgGluPheMetAspHisSerIleSer----- 253  
 Qy 747 CTGCCC---GTGGCCCGGGAGAGTTGGGTGTGGAGGACATCTTCACCGTCTCCATG 803  
 Db 254 ---ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGln 272  
 Qy 804 AATGAGAGTTTGAAGTGGTGTATATAACAGAGATGGGAAAGTGTGACCTCACGTTT 860  
 Db 273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 29  
 ADP65305  
 ID ADP65305 standard; protein; 295 AA.  
 XX  
 AC ADP65305;  
 XX  
 XX 12-AUG-2004 (first entry)  
 DT  
 XX Human sulfotransferase family, cytosolic, 1a.  
 XX  
 XX autoimmune disease; arthritis; gene expression analysis;  
 KW rheumatoid arthritis; collagen-induced; immunosuppressive; antirheumatic;  
 KW antiarthritic; osteopathic; antigout; antiinflammatory; dermatological;  
 KW immunomodulatory; lupus; ankylosing spondylitis; Fibrositis;  
 KW fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;  
 KW immune; human.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2003072827-A1.  
 PN  
 XX  
 XX 04-SEP-2003.  
 PD  
 XX  
 XX 31-OCT-2002; 2002MO-US035433.  
 PF  
 XX  
 XX 31-OCT-2001; 2001US-0336220P.  
 PR  
 XX  
 XX (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.  
 PA  
 XX  
 XX Hirsch R, Thorton SL;  
 FI  
 XX  
 XX WPI; 2003-712740/67.  
 DR  
 XX  
 XX GENBANK; NP\_001045.  
 DR  
 XX  
 XX Diagnosing and analyzing autoimmune disease using gene expression  
 PT profiles and microarray technology, useful for diagnosing and treating  
 FT rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and  
 PT gout.  
 FT  
 XX  
 XX Disclosure; Page; 56pp; English.  
 PS  
 XX  
 XX The invention relates to a novel method for diagnosing and analysing  
 CC autoimmune disease or arthritides. The method comprises obtaining a  
 CC patient sample containing mRNA, analysing gene expression using the mRNA  
 CC that results in a gene expression signature of the mRNA, and using that  
 CC gene expression signature to diagnose or analyse the autoimmune disease  
 CC or arthritides in the patient, where gene expression of at least 60% of  
 CC the genes correlates with that of the gene signature. The invention  
 CC further comprises: a treatment of rheumatoid arthritis; identification of  
 CC genes for targeting in the treatment of rheumatoid arthritis in a mammal  
 CC other than a mouse; diagnosis of rheumatoid arthritis in a mammal; an  
 CC array or gene chip, specific for rheumatoid arthritis; diagnosis or  
 CC analyses of autoimmune disease or rheumatoid arthritis; screening the  
 CC efficacy of a candidate drug in vitro for the treatment of collagen-  
 CC induced arthritis; and reducing the symptoms associated with collagen-  
 CC induced arthritis. The compositions of the invention have the following  
 CC activities: immunosuppressive, antirheumatic, antiarthritic, osteopathic,  
 CC antigout, antiinflammatory, dermatological, and immunomodulatory. The  
 CC methods and compositions of the present invention are useful for  
 CC diagnosing and treating autoimmune disease or arthritides, such as  
 CC rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis,  
 CC fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an  
 CC immune disease caused by an infectious agent. This sequence represents a

CC protein sequence relating to the genes used in the analysis and treatment  
 CC of autoimmune diseases or arthritides. Note: This sequence is not shown  
 CC in the specification. It has been supplied in an electronic format from  
 CC WIPO.  
 XX

SQ Sequence 295 AA;

Alignment Scores:  
 Pred. No.: 8.14e-34 Length: 295  
 Score: 456.00 Matches: 94  
 Percent Similarity: 54.5% Conservative: 58  
 Best Local Similarity: 33.7% Mismatches: 101  
 Query Match: 10.4% Indels: 26  
 DB: 7 Gaps: 4

US-10-768-158-1 (1-2419) x ADP65305 (1-295)

Qy 90 GGGTGGCGGTGGCGCCCTTCTGCGCGGGAAGATGGAGGATGCGCAACTTCCCGTG 149  
 Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36  
 Qy 150 CGGCCAGCAGCTGTGTGATGCTCCTACCTACCCCAAGTCCGGCACCAGCTTGTCTCGAG 209  
 Db 37 ArgProAspAspLeuLeuIleSerThrTyrProLysSerGlyThrThrTrpValSerGln 56  
 Qy 210 GTGCTCTACTTGGTGACCGCGCTGACCCGATGAGATCGCTTGTGATGAACATCGAC 269  
 Db 57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysHisArgAlaProIlePhe 76  
 Qy 270 GAGCAGCTCCCGGTCTCTGGAGTACCCACAGCCG-----GGCTGGACATCATC 317  
 Db 77 MetArgValProPheLeuGluPheLysValProGlyIleProSerGlyMetGluThrLeu 96  
 Qy 318 AAGAACTGACCTCTCCCGCTCATCAAGAGCACCCTGCTACCGCTTCTGCTCCCTCT 377  
 Db 97 LysAsnThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116  
 Qy 378 GACCTCCACAAATGGAGACTCCAAGTCTATATGCTCGCAACCCCAAGGATCTCGTG 437  
 Db 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspValAla 136  
 Qy 438 GTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGTACCGAGGCACCTTTCAA 497  
 Db 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValTyrProHisProGlyThrTrpGlu 156  
 Qy 498 GAATTCGCGGAGGTATTGAATGATAAGCTGGCTACGGCTCCTGGTGTGACAGCTG 557  
 Db 157 SerPheLeuGluLysPheMetAlaGlyValSerTyrGlySerTyrTyrGlnHisVal 176  
 Qy 558 CAGGAGTTCTGGGAGACCCGATGCGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATG 617  
 Db 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196  
 Qy 618 CATCGGACCTGTGTGACGATGGTGAGCAGCTGCCAGATTCCTGGGGGTGTCTGTGAC 677  
 Db 197 LysGluAsnProLysAspGluIleGlnLysLeuGluPheValGlyArgSerLeuPro 216  
 Qy 678 AAGGCCAGCTGGGAGCCCTGACCGACACTGC----- 710  
 Db 217 GluGluThrValAspLeuMetValGluHisThrSerPheLysGluMetLysLysAsnPro 236  
 Qy 711 -----CACAGCTGGTGACCAAGTGTGCTGCTGCAACGCTGAGGCC 746  
 Db 237 MetThrAsnTyrThrThrValArgArgGluPheMetAspHisSerIleSer----- 253  
 Qy 747 CTGCCC---GTGGCCCGGGAGAGTTGGGTGTGGAGGACATCTTCACCGTCTCCATG 803  
 Db 254 ---ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGln 272  
 Qy 804 AATGAGAGTTTGAAGTGGTGTATATAACAGAGATGGGAAAGTGTGACCTCACGTTT 860  
 Db 273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291











Db 157 SerPheLeuGluAaspPheMetAlaGlyGluValSerTyrGlySerTrpTyrGlnHisVal 176  
QY 558 CAGGAGTCTGGGAGACACCGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATG 617  
Db 177 GlnGluTrpTrpGluLeuAArgHisThrHisProValLeuTyrLeuPheTyrGluAaspMet 196  
QY 618 CATCGGAGCTGTGACGATGGTGAGACAGCTGCCAGATCTCTGGGGGTGCTGTGAC 677  
Db 197 LysGluAasnProLysAArgGluIleGlnLysIleLeuGluPheValGlyAArgSerLeuPro 216  
QY 678 AAGGCCAGCTGGAAGCCCTGACGGAGCACTGC-----CACACGCTGGTGGACCAG 728  
Db 217 GluGluThrValGluAaspIleValGlnHisThrSerPheGlnGluMetLysAsnAsnAla 236  
QY 729 TGCTGCAACGCTGAGGCCCTGCCCC-----GTG 755  
Db 237 MetThrAsnTyrArgThrLeuProSerAspLeuLeuAspHisSerIleSerAlaPheMet 256  
QY 756 GCGCGGGGAAGAGTTGGCTGTGGAGGACATCTTCACCTCTCCATGAATGAAGATTT 815  
Db 257 ArgLysGlyIleThrGlyAaspTrpLysSerThrPheThrValAlaGlnAsnGluAArgPhe 276  
QY 816 GACTTGGTGATAACAGAGATGGGAAGCTGTGACCTCAGCTT 860  
Db 277 GluAlaAaspTyrAlaGluLysMetAlaGlyCysAsnLeuAArgPhe 291  
RESULT 35  
ADD18776  
ID ADD18776 standard; protein; 295 AA.  
XX AC ADD18776;  
XX DT 15-JAN-2004 (first entry)  
XX DE Human disease related protein SeqID207.  
XX KW human; disease state; cytostatic; antiinflammatory; ophthalmological;  
KW antiarteriosclerotic; vulnery; gene therapy;  
KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;  
KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;  
KW glucose transportation; catecholamine synthesis; iron transport;  
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;  
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;  
KW inflammatory condition; wound healing.  
XX OS Homo sapiens.  
XX PN WO2003018621-A2.  
XX PD 06-MAR-2003.  
XX PF 23-AUG-2002; 2002WO-GB003892.  
XX PR 23-AUG-2001; 2001GB-00020558.  
XX PR 05-OCT-2001; 2001GB-00024037.  
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.  
XX PI Kingman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;  
XX DR N-PSDB; ADD18776.  
XX PT New substantially purified polypeptide, useful for diagnosing or treating  
XX PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion  
XX PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or  
XX PT wound healing.  
XX PS Claim 25; SEQ ID NO 207; 424pp; English.  
XX XX This invention relates to novel human genes and gene product which are  
XX XX implicated in certain disease states. Compounds which modulate the

CC proteins of the invention may have cytostatic, antiinflammatory,  
CC ophthalmological, antiarteriosclerotic or vulnery activities. The  
CC sequences of the invention may be useful for gene therapy. The invention  
CC may be useful for diagnosing or treating a hypoxia-regulated condition,  
CC such as tumorigenesis, angiogenesis, apoptosis, inflammation,  
CC erythropoiesis, or the biological response to hypoxia conditions  
CC including processes such as glycolysis, gluconeogenesis, glucose  
CC transportation, catecholamine synthesis, iron transport or nitric oxide  
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion  
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,  
CC inflammatory conditions or wound healing. The present sequence is that of  
XX a disease related protein of the invention.  
SQ Sequence 295 AA;

Alignment Scores: 2,44e-33 Length: 295  
Pred. No.: 451.00 Matches: 93  
Score: 54.5% Conservative: 57  
Best Local Similarity: 33.8% Mismatches: 107  
Query Match: 10.2% Indels: 18  
DB: Gaps: 3

US-10-768-158-1 (1-2419) x ADD18776 (1-295)

QY 90 GCGTGTGCTGCGCCCTTCTGCCGCGGAAGATGAGGAGATCGCAACTTCCCGGTG 149  
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36  
QY 150 CGGCCCGGAGCTGTGGATCGTCACCTACCCCAAGTCCGGCAGCAGCTGTGTCAGGAG 209  
Db 37 ArgProAaspLeuLeuIleSerThrTyrProLysSerGlyThrTrpValSerGln 56  
QY 210 GTGCTCTACTGTGAGCCAGCGCTGACCCCGATGATGATGCGTGTGATGACATCGAC 269  
Db 57 IleLeuAaspMetIleTyrGlnGlyAaspLeuGluLysCysHisArgAlaProIlePhe 76  
QY 270 GAGCAGCTCCCGTCTGTGAGTACCCACAGCCG-----GCCCTGGACATCATC 317  
Db 77 MetArgValProPheLeuGluPheLysAlaProGlyIleProSerGlyMetGluThrLeu 96  
QY 318 AAGGAAGTACCTCTCCCGCTCATCAAGAGCAGCTGCCCTACCGCTTCTGCTCTCT 377  
Db 97 LysAaspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116  
QY 378 GACCTCCACAATGGAGACTCCAAGTCACTATATGCTCGCAACCCCAAGGATCTGGTG 437  
Db 117 ThrLeuLeuAaspGlnLysValLysValValTyrValAlaArgAsnAlaLysAaspValAla 136  
QY 438 GTGCTTTATTCAGTTCCACCGCTCTCTGCGGACCATGAGTACCGAGGACACCTTTCAA 497  
Db 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp 156  
QY 498 GAATTCCTGCGGAGGTTTATGATGATAGTGGCTACGGCTCTGCTGTGAGCAGCTG 557  
Db 157 SerPheLeuGluLysPheMetValGlyValSerTyrGlySerTrpTyrGlnHisVal 176  
QY 558 CAGGAGTCTTGGGAGCACCGCATGGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATG 617  
Db 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAaspMet 196  
QY 618 CATCGGAGCTGTGACGATGGTGAGCAGCTGCGCAGATCTCTGGGGGTGCTGTGAC 677  
Db 197 LysGluAasnProLysAArgGluIleGlnLysIleLeuGluPheValGlyHisSerLeuPro 216  
QY 678 AAGGCCAGCTGGAAGCCCTGACGGAGCACTGC-----CACACGCTGGTGGACCAG 728  
Db 217 GluGluThrValAaspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236  
QY 729 TGCTGCAACGCTGAGGCCCTGCCCC-----GTG 755  
Db 237 MetThrAsnTyrThrThrValProGlnGluPheMetAspHisSerIleSerProPheMet 256

QY 756 GGCCGGGAGAGTGGCGTGTGGNAGGACATCTTACCGCTCCCATGAATGAGAGTTT 815  
Db : : : : :  
257 ArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGlnAenGluArgPhe 276  
QY 816 GACTTGGTGATAAACAAGAGATGGGAAAGTGTGACCTCACGTTT 860  
Db : : : : :  
277 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291  
RESULT 36  
ID ADE57149  
AD ADE57149 standard; protein; 295 AA.  
XX  
AC ADE57149;  
XX  
XX  
DT 29-JAN-2004 (first entry)  
DE Human Protein P50225, SEQ ID NO 3009.  
XX  
XX Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
XX  
XX Homo sapiens.  
XX  
XX W02003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
XX 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
XX (GEHO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
XX Woolf C, D'urso D, Befort K, Costigan M;  
XX  
XX WPI; 2003-268312/26.  
DR GENBANK; P50225.  
XX  
XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
XX Claim 1; Page; 1017pp; English.  
XX  
XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative of allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 295 AA;  
Alignment Scores: 2.44e-33 Length: 295  
Pred. No.: 451.00 Matches: 93  
Score: 54.5% Conservativeness: 57  
Percent Similarity: 33.8% Mismatches: 107  
Best Local Similarity: 10.2% Indels: 18  
Query Match: 7 Gaps: 3  
DB:  
US-10-768-158-1 (1-2419) x ADE57149 (1-295)  
QY 90 GCGTGGCGTCCGCCCTTCTGCCGGGAGATGGAGGAGATGCCAACACTTCCCGGTG 149  
Db : : : : :  
17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36  
QY 150 CGGCCCGAGCGTGTGGATCGTCACCTACCCCAAGTCCGGCACCAGCTTGTGTCAGGAG 209  
Db : : : : :  
37 ArgProAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpValSerGln 56  
QY 210 GTGGTCTACTTGTGAGCCAGCGCTGACCCCGATCGATCGCTTGTGAACATCGAC 269  
Db : : : : :  
57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysHisArgAlaProIlePhe 76  
QY 270 GAGCAGCTCCCGTCTCGAGTACCCACAGCCG-----GGCTGGACATCATC 317  
Db : : : : :  
77 MetArgValProPheLeuGluPheLysAlaProGlyIleProSerGlyMetGluThrLeu 96  
QY 318 AAGGAAGTACCTCTCCCGCTCATCAAGAGCCACTGCCCTACCGCTTCTGCCCTCT 377  
Db : : : : :  
97 LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116  
QY 378 GACCTCCACAATGGAGACTCCAAGGTCTATATATGCTCGCAACCCCAAGGATCTGGTG 437  
Db : : : : :  
117 ThrLeuLeuAspGlnLysValValValValValValValValValValValValVal 136  
QY 438 GTGCTTTATATCATGTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGGCACCTTCAA 497  
Db : : : : :  
137 ValSerTyrTyrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp 156  
QY 498 GAATTTCTGCGGAGGTTTATGAATAGCTAGCTGGCTACGGCTCTGTTTGGACGAGTG 557  
Db : : : : :  
157 SerPheLeuGluLysPheMetValGlyGluValSerTyrGlySerTyrGlnHisVal 176  
QY 558 CAGGAGTCTTGGGAGCACCAGCATGAGCTCGAAGCTGCTTTTCTCAAGTATGAAGCATG 617  
Db : : : : :  
177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196  
QY 618 CATCGGACCTGTGAGTGGATGGTGGAGCAGCTGCGCAGATTCCTGGGGTGTCTGTGAC 677  
Db : : : : :  
197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216  
QY 678 AAGGCCCGAGCTGAAGCCCTGACGGAGCAGTCG-----CACCAGCTGGTGGACAG 728  
Db : : : : :  
217 GlnGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236  
QY 729 TGCTCAACGCTGAGGCCCTGCCCC-----GTG 755  
Db : : : : :  
237 MetThrAsnTyrThrThrValProGlnGluPheMetAspHisSerIleSerProPheMet 256  
QY 756 GCGCGGGGAGAGTGGGTGTGGAGGACATCTTACCGCTCCCATGAATGAGAGTTT 815  
Db : : : : :  
257 ArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGlnAenGluArgPhe 276  
QY 816 GACTTGGTGATAAACAAGAGATGGGAAAGTGTGACCTCACGTTT 860  
Db : : : : :  
277 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291  
RESULT 37  
ADE57137  
ID ADE57137 standard; protein; 295 AA.

XX	AC	ADBS57137;	US-10-768-158-1 (1-2419) x ADBS57137 (1-295)
XX	DT	29-JAN-2004 (first entry)	Qy 90 GGCCTGCGGCTGCGCGCTCTCTGCGCGGGAAGATGGAGGAGATCGCCAACTTCCCGGTG 149
XX	DE	Human Protein P50225, SEQ ID NO 2997.	Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
XX	KW	Human; pain; neuronal tissue; gene therapy;	Qy 150 CGGCCAGGACGACGTGTGGATCGTCACCTACCCCAAGTCCCGCACCAGCTTCTCGAGGAG 209
XX	KW	spinal segmental nerve injury; chronic constriction injury; CCI;	Db 37 ArgProAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpLysGln 56
XX	OS	Homo sapiens.	Qy 210 GTGCTCTACTTGGTGAGCCGCGCTGACCCCGATGAGATCGGCTTCATCAACATCCAC 269
XX	PN	WO2003016475-A2.	Db 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGlnLysCysHisArgAlaProIlePhe 76
XX	PD	27-FEB-2003.	Qy 270 GAGCAGCTCCCGGCTCTGGAGTACCCACAGCCG-----GGCCTGGACATCATC 317
XX	PF	14-AUG-2002; 2002WO-US025765.	Db 77 MetArgValProPheLeuGluPheLysAlaProGlyIleProSerGlyMetGluThrLeu 96
XX	PR	14-AUG-2001; 2001US-0312147P.	Qy 318 AAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCGCTTCTGCTTCTGCTTCT 377
XX	PR	01-NOV-2001; 2001US-0346382P.	Db 97 LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116
XX	PA	(GEHO ) GEN HOSPITAL CORP.	Qy 378 GACCTCCACAATGGAGACTCCAGGTCTATATGCTCGCACCCCAAGATCTGGTG 437
XX	PA	(FARB ) BAYER AG.	Db 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspValAla 136
XX	PI	Woolf C, D'urso D, Befort K, Costigan M;	Qy 438 GTGCTTTATTATCAGTTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGGACCTTCAA 497
XX	DR	WPI; 2003-268312/26.	Db 137 ValSerTyrThrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp 156
XX	DR	GENBANK; P50225.	Qy 498 GAATTCGCGGAGGTTTATGAATGATAGCTGGGCTACCGCTCCTGTTTGGACAGCTG 557
XX	PT	New composition comprising two or more isolated polypeptides, useful for	Db 157 SerPheLeuGluLysPheMetValGlyLysValSerTyrGlySerTyrTyrGlnHisVal 176
XX	XX	preparing a medicament for treating pain in an animal.	Qy 558 CAGGAGTCTGGGAGCACCATGAGTCTCGAACGTCGCTTTTCTCAAGTATGAGACATG 617
XX	PS	Claim 1; Page; 1017pp; English.	Db 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
XX	CC	The invention discloses a composition comprising two or more isolated rat	Qy 618 CATCGGACCTGGTGACGATGGTGAGCAGCTGGCCAGATTCCTGGGGGTCTCTGTGAC 677
XX	CC	or human polynucleotides or a polynucleotide which represents a fragment,	Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
XX	CC	derivative or allelic variation of the nucleic acid sequence. Also	Qy 678 AAGCCCGAGCTGGAAGCCCTGACGAGCAGCTGC-----CACCACTGGTGGACAG 728
XX	CC	claimed are a vector comprising the novel polynucleotide, a host cell	Db 217 GluGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236
XX	CC	comprising the vector, a method for identifying a nucleotide sequence	Qy 729 TGCTGCAACGCTGAGGCCCTGCCCC-----GTG 755
XX	CC	which is differentially regulated in an animal subjected to pain and a	Db 237 MetThrAsnTyrThrThrValProGlnGluPheMetAspHisSerIleSerProPheMet 256
XX	CC	kit to perform the method, an array, a method for identifying an agent	Qy 756 GCGCGGGAAGAGTTGGGCTGTGGAAGGACATCTTACCGCTCTCCATGAATGAGAAGTTT 815
XX	CC	that increases or decreases the expression of the polynucleotide sequence	Db 257 ArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGlnAsnGluArgPhe 276
XX	CC	that is differentially expressed in neuronal tissue of a first animal	Qy 816 GACTTGGTGTATAACAGAGATGGAAAGTGTGACCTCAGCTTT 860
XX	CC	subjected to pain, a method for identifying a compound which regulates	Db 277 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291
XX	CC	the expression of a polynucleotide sequence which is differentially	RESULT 38
XX	CC	expressed in an animal subjected to pain, a method for identifying a	ADBS57141
XX	CC	compound that regulates the activity of one or more of the	ID ADBS57141 standard; protein; 295 AA.
XX	CC	polynucleotides, a method for producing a pharmaceutical composition, a	XX AC ADBS57141;
XX	CC	method for identifying a compound or small molecule that regulates the	XX DT 29-JAN-2004 (first entry)
XX	CC	activity in an animal of one or more of the polypeptides given in the	XX DE Human Protein P50225, SEQ ID NO 3001.
XX	CC	specification, a method for identifying a compound useful in treating	XX KW Human; pain; neuronal tissue; gene therapy;
XX	CC	pain and a pharmaceutical composition comprising the one or more	XX KW spinal segmental nerve injury; chronic constriction injury; CCI;
XX	CC	polypeptides or their antibodies. The polynucleotide or the compound that	XX KW spared nerve injury; SNI; Chung.
XX	CC	modulates its activity is useful for preparing a medicament for treating	XX DB:
XX	CC	pain (e.g. spinal segmental nerve injury (Chung), chronic constriction	
XX	CC	injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene	
XX	CC	therapy). The sequence presented is a human protein (shown in Table 2 of	
XX	CC	the specification) which is differentially expressed during pain. Note:	
XX	CC	The sequence data for this patent did not form part of the printed	
XX	CC	specification, but was obtained in electronic form directly from WIPO at	
XX	CC	ftp.wipo.int/pub/published_pct_sequences.	
XX	SQ	Sequence 295 AA;	
XX	Alignment Scores:		
XX	Pred. No.:	2.44e-33	Length: 295
XX	Score:	451.00	Matches: 93
XX	Percent Similarity:	54.5%	Conservative: 57
XX	Best Local Similarity:	33.8%	Mismatches: 107
XX	Query Match:	10.2%	Indels: 18
XX	DB:	7	Gaps: 3







XX PS Claim 2; SEQ ID NO 245; 323pp; Japanese.

XX CC The invention relates to a tumour antigen peptide recognised by human

CC leukocyte antigen (HLA)-A2 or HLA-A26 restricted cytotoxic T-cells (CTL)

CC and/or capable of inducing CTL. The tumour antigen peptide is useful for

CC the treatment, prevention, diagnosis and vaccine production for cancers

CC including colorectal, stomach, buccal, renal, lung, gynecological and

CC prostate cancer. The present sequence represents the amino acid sequence

CC of a human protein.

XX SQ Sequence 295 AA;

Alignment Scores:

Pred. NO.:	2.44e-33	Length:	295
Score:	451.00	Matches:	93
Percent Similarity:	54.5%	Conservative:	57
Best Local Similarity:	33.8%	Mismatches:	107
Query Match:	10.2%	Indels:	18
DB:	7	Gaps:	3

US-10-768-158-1 (1-2419) x ADI15901 (1-295)

QY 90 GGCCTGGCGCTCCGCGCTCTCTGCGGGGAGATGGAGAGATCGCAACTTCCCGTG 149

Db 17 GlyValProLeuIleLeuTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36

QY 150 CGGCCCGCAGCGAGTGTGGATCGTCACTTACCCCAAGTCCGGCACCAGCTTGTCTGCAGGAG 209

Db 37 ArgProAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpValSerGln 56

QY 210 GTGCTCTACTTGGTGAGCCGGCGCTGACCCCATGAGATCGCGCTTGATGAACATCGAC 269

Db 57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysHisArgAlaProIlePhe 76

QY 270 GAGCAGCTCCCGCTCTGGAGTACCACAGCCG-----GGCTCGACATCATC 317

Db 77 MetArgValProPheLeuGluPheLysAlaProGlyIleProSerGlyMetGluThrLeu 96

QY 318 AAGAACTGACCTCTCCCGCTCATCAAGAGCCAGCTGCTTCTCCCTCT 377

Db 97 LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuProGln 116

QY 378 GACTCCACATGAGACTCCAGGTCTATGCTGCTGCAACCCAGGATCTGGTG 437

Db 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspVala 136

QY 438 GTGCTTATTATCAGTTCACCGCTCTCTGCGGACCATGAGCTACCGAGGCACCTTCA 497

Db 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp 156

QY 498 GAATTTCTCCGAGGTTTATGAATGATAGCTGGGTACGCTCTGTTTGGACAGTG 557

Db 157 SerPheLeuGluLysPheMetValGlyLysValSerTyrGlySerTyrTyrGlnHisVal 176

QY 558 CAGAGTTCTGGAGCACCAGCATGAGCTGCAACGTGCTTTTCTCAGATGAGACATG 617

Db 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196

QY 618 CATCGGACCTGGTGACGATGGTGAGCAGCTGGCGAGATTCCTGGGGGTCTCTGTGAC 677

Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyHisSerLeuPro 216

QY 678 AAGCCCGCAGCTGGAAGCCCTGACGAGCATGTC-----CACCACTGGTGACCAAG 728

Db 217 GluGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236

QY 729 TGCTGCACGCTAGGCCCTGCC-----GTG 755

Db 237 MetThrAsnTyrThrThrValProGlnGluPheMetAspHisSerIleSerProPheMet 256

QY 756 GCGCGGGAAGAGTTGGGCTGTGGAAGGACATCTTCCCGTCTCCATGAATGAGAAGTTT 815

Db 257 ArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGlnAsnGluArgPhe 276

QY 816 GACTTGGTGTATTAACAGAGATGGGAAAGTGTGACCTCAGTTT 860

Db 277 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 41

ADP23588

ID ADP23588 standard; protein; 295 AA.

XX AC ADP23588;

XX 18-NOV-2004 (first entry)

XX PRO polypeptide SEQ ID NO:766.

DE PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;

KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;

KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.

XX Unidentified.

OS WO2004041170-A2.

PN 21-MAY-2004.

PD 30-OCT-2003; 2003WO-US034312.

PF 01-NOV-2002; 2002US-0423394P.

PR (GETH ) GENENTECH INC.

PA Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;

PI Wu TD;

XX WPI; 2004-419628/39.

DR N-PSDB; ADP23587.

XX New PRO polypeptides and polynucleotides, useful for treating e.g.

PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated

PT renal disease, or demyelinating diseases of the central or peripheral

PT nervous system.

XX Claim 7; SEQ ID NO 766; 2940pp; English.

PS The invention relates to a novel isolated nucleic acid and the PRO

XX polypeptide encoded by it. A protein of the invention has

CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,

CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,

CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide

CC of the invention may have a use in gene therapy. The PRO polypeptide, its

CC agonist, antagonist, or antibody that specifically binds to the

CC polypeptide is useful for treating an immune related disorder such as

CC systemic lupus erythematous, rheumatoid arthritis, osteoarthritis,

CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an

CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic

CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune

CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal

CC disease, a demyelinating disease of the central or peripheral nervous

CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,

CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary

CC disease, infectious or autoimmune chronic active hepatitis, primary

CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,

CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's

CC disease, an autoimmune or immune-mediated skin disease, a bullous skin

CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic

CC disease, asthma, allergic rhinitis, atopic dermatitis, food

CC hypersensitivity, urticaria, an immunologic disease of the lung,

CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity

CC pneumonitis, a transplantation associated disease, graft rejection or

CC graft-versus-host disease. The present sequence represents a PRO protein

XX of the invention.

SQ	Sequence 295 AA;			
Alignment Scores:				
Pred. No.:	2.44e-33	Length:	295	
Score:	451.00	Matches:	93	
Percent Similarity:	54.5%	Conservative:	57	
Best Local Similarity:	33.8%	Mismatches:	107	
Query Match:	10.2%	Indels:	18	
DB:	8	Gaps:	3	
US-10-768-158-1 (1-2419) x ADP23588 (1-295)				
Qy	90	GGCGTGGCGTGGCGCCCTTCTGCGCGGGAAGATGGAGGATCGCCAACTTCCCGGTG	149	
Db	17	GlyValProLeuIleLysTyPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla	36	
Qy	150	CGGCCCGGAGCGTGTGTGATCGTACCTACCCCAAGTCGGCAGCCAGCTGCTGCAGGAG	209	
Db	37	ArgProAspLeuLeuIleSerThrTyProLysSerGlyThrTrpValSerGln	56	
Qy	210	GTGCTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC	269	
Db	57	IleLeuAspMetIleTyGlnGlyAspLeuGluLysCysHisArgAlaProIlePhe	76	
Qy	270	GAGCAGCTCCCGTCTGGAGTACCCACAGCGG-----GGCTGGACATCATC	317	
Db	77	MetArgValProPheLeuGluPheLysAlaProGlyIleProSerGlyMetGluThrLeu	96	
Qy	318	AAGGAACGTGACTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTCTGCCCTCT	377	
Db	97	LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln	116	
Qy	378	GACCTCCCAATGGAGACTCAAGGTCTATATGGCTCGCAACCCCAAGATCTGGTG	437	
Db	117	ThrLeuLeuAspGlnLysValLysValValTyValAlaAaGAsnAlaLysAspValAla	136	
Qy	438	GTGCTTATTATCATCTCCACCGTCTCTCGGACCATGAGCTACCGAGGACCTTTCAA	497	
Db	137	ValSerTyTyHisPheTyHisMetAlaLysValHisProGluProGlyThrTrpAsp	156	
Qy	498	GAATCTCCCGAGGTTTATGAATGATAAGCTGAGCTACGCTCTGTTTGGACACGTG	557	
Db	157	SerPheLeuGluLysPheMetValGlyGluValSerTyGlySerTrpTyGlnHisVal	176	
Qy	558	CAGGAGTTCTGGGACACCGCATGCACTGCAACGTGCTTTTCTCAAGTATGAAGACATG	617	
Db	177	GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyLeuPheTyGluAspMet	196	
Qy	618	CATCGGACCTGGTGACGATGGTGAGCAGCTGCGCAGATTCCTGGGGTGTCTGTGAC	677	
Db	197	LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyHisSerLeuPro	216	
Qy	678	AAGGCCAGCTGGAAGCCCTGACGAGCACTGC-----CACCAGCTGGTGGACACAG	728	
Db	217	GluGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro	236	
Qy	729	TGCTCAACGCTGAGCGCTGCCC-----GTG	755	
Db	237	MetThrAsnTyThrThrValProGlnGluPheMetAspHisSerIleSerProPheMet	256	
Qy	756	GGCCGGGAAGAGTTGGCTGTGGGAAGACATCTTCACCGTCTCCATGAATGAGAAGTTT	815	
Db	257	ArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPhe	276	
Qy	816	GACTTGGTGTTAAACAGAAAGTGGAAAGTGTGACCTCAGCTT	860	
Db	277	AspAlaAspTyAlaGluLysMetAlaGlyCysSerLeuSerPhe	291	
RESULT 42				
ID	ADR97362	standard; protein; 295 AA.		
XX				
AC	ADR97362;			

XX	02-DEC-2004	(first entry)	
DT	Human SULT1A1	protein, an apoptosis related target	Seq 70.
XX	human; enzyme; apoptosis; cancer; inflammation; autoimmune; neurodegenerative disorder; cytostatic; antiinflammatory; immunosuppressive; neuroprotective; gene therapy; SULT1A1; sulfotransferase family.		
XX	Homo sapiens.		
XX	WO2004078783-A2.		
XX	16-SEP-2004.		
XX	05-MAR-2004; 2004WO-GB000957.		
XX	07-MAR-2003; 2003GB-00005267.		
XX	(EIRX-) EIRX THERAPEUTICS LTD.		
XX	Murphy FJ, Sheehan DE, Keating KE, Hayes I, Seera L;		
XX	WPI; 2004-662402/64.		
XX	N-PSDB; ADR97361.		
PT	Identifying an agent that modulates the function of an apoptosis-associated polypeptide, useful for diagnosing or treating e.g. cancer, comprises comparing the binding of the polypeptide to the candidate agent and to a control agent.		
XX	Claim 1; SEQ ID NO 70; 304pp; English.		
CC	This invention relates to novel agents that modulates the function of human apoptosis-associated proteins specified within the specification. Specifically, it refers to a method for the identification of target genes whose expression is correlated with an early stage in the regulation of apoptosis. The present invention describes a method of contacting either candidate agents or control agents to the target genes and assessing the difference of binding and inhibitory activity, where the candidate agent is selected from a low molecular weight organic molecule, an antibody or fragment thereof, an antisense oligonucleotide, a small inhibitory dsRNA, or a ribozyme. As such, the compositions and methods are useful for diagnosing and treating diseases or conditions associated with abnormal apoptosis in mammalian tissue, such as cancer, inflammation, autoimmune or neurodegenerative disorders. Accordingly, they exhibit cytostatic, antiinflammatory, immunosuppressive and neuroprotective activities. These may also be used for drug screening purposes and in gene therapy. This polypeptide sequence is a human target protein, an enzyme associated with the regulation of apoptosis whose expression is modulated by novel agents of the invention.		
SQ	Sequence 295 AA;		

Alignment Scores:			
Pred. No.:	2.44e-33	Length:	295
Score:	451.00	Matches:	93
Percent Similarity:	54.5%	Conservative:	57
Best Local Similarity:	33.8%	Mismatches:	107
Query Match:	10.2%	Indels:	18
DB:	8	Gaps:	3

US-10-768-158-1 (1-2419) x ADR97362 (1-295)

Qy	90	GGCGTGGCTGCGCGCCCTTCTGCGCGGGAAGATGGAGGATCGCCAACTTCCCGTG	149
Db	17	GlyValProLeuIleLysTyPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla	36
Qy	150	CGGCCCGGAGCGTGTGTGATCGTACCTACCCCAAGTCGGCAGCCAGCTTGTCTCAGGAG	209
Db	37	ArgProAspLeuLeuIleSerThrTyProLysSerGlyThrTrpValSerGln	56





Score:	449.00	Matches:	96
Percent Similarity:	55.8%	Conservative:	52
Best Local Similarity:	36.2%	Mismatches:	99
Query Match:	10.2%	Indels:	18
DB:	5	Gaps:	3

US-10-768-158-1 (1-2419) x ADI17050 (1-302)

QY	126	GAGAGATCGCAACTTC	CGCGTCCGCGCCAGCGACGCTGTGGATCGTCACTACCCCAAG	185
DB	36	AspIyleItrpshnPh	eGlnAlaLysProAspAspLeuLeuIleSerThrTyrProLys	55
QY	186	TCCGGCACCAGCTTGCT	GAGAGGTGGTCTACTTGGTGAGCCAGGGCGCTGACCCCGAT	245
DB	56	AlaGlyThrThrTrp	ThrGlnGluIleValGluLeuIleGlnAsnGluGlyAspValGlu	75
QY	246	GAGATCGGTTGATGAAC	ATCGAGCAGCAGCTCCCGTCTCGAGTACCCACGCG---	302
DB	76	LysSerLysArgAla	ProThrHisGlnArgPheProPheLeuGluMetLysIleProSer	95
QY	303	-----GGCTGGACAT	CATCAAGGAAGTACCTCCCGCTCATCAAGGCCAC	353
DB	96	LeuGlySerGlyLeu	GluGlnAlaHisAlaMetProSerProArgIleLeuLysThrHis	115
QY	354	CTGCCCTACCGCTTT	CTGCCCTCGACCTCCCAATGGAGACTCCCAAGTCACTATATG	413
DB	116	LeuProPheHisLeu	LeuProProSerLeuLeuGluLysAsnCysLysIleIleTyrVal	135
QY	414	GCTCGCAACCCCAAG	ATCTGGTGGTCTTATATCATGTTCCACGGCTCTCTCGCGACC	473
DB	136	AlaArgAsnProLys	AspAsnMetValSerTyrTyrHisPheGlnArgMetAsnLysAla	155
QY	474	ATGAGCTACCGAGC	ACCTTCAAGATTCTCGCGGAGGTTTATGAATAGTCTGGC	533
DB	156	LeuProAlaProGly	ThrTrpGluGluTyrPheGluThrPheLeuAlaGlyLysValCys	175
QY	534	TACGGCTCTCGTTT	GAGACGTCGAGGAGTCTCGGAGCAGCGCTGCAACGCTG	593
DB	176	TrpGlySerTrpHis	GluHisValLysGlyTrpTrpGluAlaLysAspLysHisArgIle	195
QY	594	CTTTTCTCAAGTAT	GAAGCATGCATCGGACCTGGTGACGATGGTGAGCAGCTGGCC	653
DB	196	LeuTyrLeuPheTyr	GluAspMetLysAsnProLysHisGluIleGlnLysLeuAla	215
QY	654	AGATTCTGGGGTG	TCTGTGACAGGCCAGCTCGGAAGCCCTGACGAGCAGCTGCCAC	713
DB	216	GluPheIleGlyLys	LysLeuAspLysValLeuAspLysIleValHisTyrThrSer	235
QY	714	CAGCTGGTG-----	GACCAAGTCTGCAACGCTGAGCGCTCGCC	752
DB	236	PheAspValMetLys	GlnAsnProMetAlaAsnTyrSerSerIleProAlaGluIleMet	255
QY	753	-----GTGGCGGG	GAAGATTGGGCTGTGGAGGACATCTTC	791
DB	256	AspHisSerIleSer	ProPheMetArgLysGlyAlaValGlyAspTrpLysLysHisPhe	275
QY	792	ACGCTCTCCATGAT	GAGAAGTTTGCTGGTGATATAACAGAGTGGGAAGTGTGAC	851
DB	276	ThrValAlaGlnAsn	GluArgPheAspGluIleAspTyrLysLysLysMetThrAspThrArg	295
QY	852	CTCAGCTTTGACTTT	866	
DB	296	LeuThrPheHisPhe	300	

RESULT 45  
ADP76915  
ID ADP76915 standard; protein; 302 AA.  
XX AC ADP76915;  
XX 26-FEB-2004 (first entry)  
DT Novel human secreted and transmembrane protein SeqID 590.  
DE

human; PRO; membrane bound protein; membrane bound receptor; cell proliferation; cell migration; cell differentiation; mitogenic factor; survival factor; cytotoxic factor; differentiation factor; neuropeptide; hormone; cell receptor; receptor-ligand interaction; cytosolic; chondrocyte; tumour.  
Homo sapiens.  
W02003072035-A2.  
04-SEP-2003.  
21-FEB-2003; 2003WO-US005241.  
22-FEB-2002; 2002US-0359461P.  
(GETH ) GENENTECH INC.  
Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR; Williams PM, Wood WI, Wu TD;  
WPI; 2003-721702/68.  
N-PSDB; ADF76914.  
New PRO polypeptides, useful for diagnosing and treating an immune related disorder, e.g. systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or diabetes mellitus.  
Claim 10; SEQ ID NO 590; 918pp; English.  
This invention relates to novel nucleic acids encoding human PRO secreted and transmembrane proteins. Extracellular proteins play important roles in the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells (for example proliferation, migration or differentiation) is typically governed by information received from other cells and the immediate environment. The information is often transmitted by secreted polypeptides (for example mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides or hormones) which are received and interpreted by diverse cell receptors or membrane bound proteins. These membrane bound proteins and receptors may be of use as pharmaceutical and diagnostic agents, such as in the blocking of receptor-ligand interactions. The current invention provides the amino acid sequences of novel human membrane bound receptors and proteins, along with the cDNA sequences encoding them. The novel proteins of the invention may have cytosolic activities through the stimulation of chondrocytes. The nucleic acids of the invention may be useful for the manufacture of a medicament for diagnosing or treating a tumour in a mammal. In addition, they may be useful for measuring or detecting the expression of a tumour associated gene. The present sequence is the amino acid sequence of a human PRO protein of the invention.

Sequence 302 AA;

Alignment Scores:

Pred. No.:	3,83e-33	Length:	302
Score:	449.00	Matches:	96
Percent Similarity:	55.8%	Conservative:	52
Best Local Similarity:	36.2%	Mismatches:	99
Query Match:	10.2%	Indels:	18
DB:	7	Gaps:	3

US-10-768-158-1 (1-2419) x ADF76915 (1-302)

QY	126	GAGAGATCGCAACTTC	CGCGTCCGCGCCAGCGACGCTGTGGATCGTCACTACCCCAAG	185
DB	36	AspIyleItrpshnPh	eGlnAlaLysProAspAspLeuLeuIleSerThrTyrProLys	55
QY	186	TCCGGCACCAGCTTGCT	GAGAGGTGGTCTACTTGGTGAGCCAGGGCGCTGACCCCGAT	245
DB	56	AlaGlyThrThrTrp	ThrGlnGluIleValGluLeuIleGlnAsnGluGlyAspValGlu	75
QY	246	GAGATCGGTTGATGAAC	ATCGAGCAGCAGCTCCCGTCTCGAGTACCCACGCG---	302
DB	76	LysSerLysArgAla	ProThrHisGlnArgPheProPheLeuGluMetLysIleProSer	95
QY	303	-----GGCTGGACAT	CATCAAGGAAGTACCTCCCGCTCATCAAGGCCAC	353
DB	96	LeuGlySerGlyLeu	GluGlnAlaHisAlaMetProSerProArgIleLeuLysThrHis	115
QY	354	CTGCCCTACCGCTTT	CTGCCCTCGACCTCCCAATGGAGACTCCCAAGTCACTATATG	413
DB	116	LeuProPheHisLeu	LeuProProSerLeuLeuGluLysAsnCysLysIleIleTyrVal	135
QY	414	GCTCGCAACCCCAAG	ATCTGGTGGTCTTATATCATGTTCCACGGCTCTCTCGCGACC	473
DB	136	AlaArgAsnProLys	AspAsnMetValSerTyrTyrHisPheGlnArgMetAsnLysAla	155
QY	474	ATGAGCTACCGAGC	ACCTTCAAGATTCTCGCGGAGGTTTATGAATAGTCTGGC	533
DB	156	LeuProAlaProGly	ThrTrpGluGluTyrPheGluThrPheLeuAlaGlyLysValCys	175
QY	534	TACGGCTCTCGTTT	GAGACGTCGAGGAGTCTCGGAGCAGCGCTGCAACGCTG	593
DB				

```
Qy 246 GAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGTCTGGAGTACCCACAGCCG--- 302
   ::
Db 76 LysSerLysArgAlaProThrHisGlnArgPheProPheLeuGluMetLysIleProSer 95
   ::::: ||| |||||
Qy 303 -----GGCTGGACATCATCAGAACTGACCTCCCGCTCATCAAGAGCCAC 353
   ::::: ||| ||||| |||||
Db 96 LeuGlySerGlyLeuGluGlnAlaHisAlaMetProSerProArgIleLeuLysThrHis 115
   ::::: ||| ||||| |||||
Qy 354 CTGCCCTACCGCTTCTCCCTCTGACCTCCACAAATGGAGACTCCAAAGGTTCATCTATATG 413
   ::::: ||||| ||||| |||||
Db 116 LeuProPheHisLeuLeuProSerLeuLeuGluLysAsnCysLysIleIleTyrVal 135
   ::::: ||||| ||||| |||||
Qy 414 GCTCGCAACCCCAAGGATCTGGTGTGCTTATTATCATGTTCCACCGCTCTCTCGGACC 473
   ::::: ||||| ||||| |||||
Db 136 AlaArgAsnProLysAspAsnMetValSerTyrHisPheGlnArgMetAsnLysAla 155
   ::::: ||||| ||||| |||||
Qy 474 ATGAGCTACCGAGCAGCCTTTCAAGAACTTCGCGGAGGTTTATGAATGATAAGCTGGC 533
   ::::: ||||| ||||| |||||
Db 156 LeuProAlaProGlyThrTrpGluGluTyrPheGluThrPheLeuAlaGlyLysValCys 175
   ::::: ||||| ||||| |||||
Qy 534 TACGGCTCTCGTGTGAGCAGCTGCAGGAGTTCGGGAGCACCGCATGGACTCGAAAGCTG 593
   ::::: ||||| ||||| |||||
Db 176 TrpGlySerTrpHisGluHisValLysGlyTrpTrpGluAlaLysAspLysHisArgIle 195
   ::::: ||||| ||||| |||||
Qy 594 CTTTCTTCAAGTATGAAGACATCGGACCTGCTGTGACGATGGTGAGAGCTGGCC 653
   ::::: ||||| ||||| |||||
Db 196 LeuTyrLeuPheTyrGluAspMetLysLysAsnProLysHisGluIleGlnLysLeuAla 215
   ::::: ||||| ||||| |||||
Qy 654 AGATTCTCGGGGTCTCTGTGACAGCCCGAGCTGGAGCCCTGAGGACCATGCGCAC 713
   ::::: ||||| ||||| |||||
Db 216 GluPheIleGlyLysLysLeuAspLysValLeuAspLysIleValHisTyrThrSer 235
   ::::: ||||| ||||| |||||
Qy 714 CAGCTGTGTG-----GACCAGTGTGCAACGCTGAGGCGCTGCC----- 752
   ::::: ||||| ||||| |||||
Db 236 PheAspValMetLysGlnAsnProMetAlaAsnTyrSerSerIleProAlaGluIleMet 255
   ::::: ||||| ||||| |||||
Qy 753 -----GTGGCGCGGGAAGAGTTGGCTGTGGAGGACATCTTC 791
   ::::: ||||| ||||| |||||
Db 256 AspHisSerIleSerProPheMetArgLysGlyAlaValGlyAspTrpLysLysHisPhe 275
   ::::: ||||| ||||| |||||
Qy 792 ACCGTCTCCATGATGAGAGTTTGCATTGTGTATATAACAGAGATGGGAAGTGTGAC 851
   ::::: ||||| ||||| |||||
Db 276 ThrValAlaGlnAsnGluArgPheAspGluAspTyrLysLysMetThrAspThrArg 295
   ::::: ||||| ||||| |||||
Qy 852 CTCACGTTTGACTTT 866
   ::::: ||||| ||||| |||||
Db 296 LeuThrPheHisPhe 300
   ::::: ||||| ||||| |||||
RESULT 46
ADZ70333
ID ADZ70333 standard; protein; 302 AA.
XX
XX
AC ADZ70333;
XX
DT 30-JUN-2005 (first entry)
XX
DE Human protein from lung cancer marker gene SUL1C2.
XX
XX Tumor marker; lung tumor; cytostatic; neoplasm; expression;
KW DNA microarray.
XX
OS Homo sapiens.
XX
PN W02005032495-A2.
XX
XX 14-APR-2005.
PD
XX
XX 01-OCT-2004; 2004WO-US034163.
XX
XX 03-OCT-2003; 2003US-0508355P.
XX
XX (FARB ) BAYER PHARM CORP.
PA
```

```
XX
PI Taylor I, Pauloski NR, Bigwood D;
XX WPI: 2005-285325/29.
DR N-PSDB; ADZ70332.
XX
XX Providing a patient diagnosis for lung cancer comprises comparing the
FT level of expression of genes or gene products in a biological sample from
FT the patient with that from a normal individual.
XX
XX Claim 3; SEQ ID NO 18; 60pp; English.
XX
XX The invention relates to providing a patient diagnosis for lung cancer
CC comprising comparing the level of expression of genes or gene products in
CC a biological sample from the patient with the level of expression of
CC genes or gene products in a biological sample from a normal individual.
CC Also included are distinguishing between normal and disease tissues,
CC monitoring the response of a patient being treated for lung cancer by
CC administering an anti-cancer agent, identifying a compound useful for the
CC treatment of lung cancer and an array for distinguishing between normal
CC and disease tissues (comprising 2 or more probes corresponding to 2 or
CC more genes selected from any of the 200 nucleotide sequences given in the
CC specification, or 2 or more polypeptides comprising any of the 200 amino
CC acid sequences given in the specification). In providing a patient
CC diagnosis for lung cancer, one or more genes are selected from any of the
CC 200 nucleotide sequences as mentioned in the specification, or one or
CC more gene products are polypeptides selected from any of the 20 amino
CC acid sequences mentioned in the specification. The methods are useful for
CC detecting and treating lung cancer. These may also be used for designing,
CC identifying and optimizing therapeutics for cancer. The present sequence
CC represents a protein from one of the 200 lung cancer marker genes. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 302 AA;
Alignment Scores:
Pred. No.: 3,83e-33 Length: 302
Score: 449.00 Matches: 96
Percent Similarity: 55.8% Conservative: 52
Best Local Similarity: 36.2% Mismatches: 99
Query Match: 10.2% Indels: 18
DB: Gaps: 3
US-10-768-158-1 (1-2419) x ADZ70333 (1-302)
Qy 126 GAGGAGATCGCAACTTCCCGTCCGCGCAGGACGCTGTGGATCGTCACCTACCCCAAG 185
   ::::: ||||| ||||| |||||
Db 36 AspLysIleTrpAsnPheGlnAlaLysProAspLeuLeuIleSerThrTyrProLys 55
   ::::: ||||| ||||| |||||
Qy 186 TCCGGCACCACTTGCTGTCAGGAGGTGCTACTTGTGTGAGCCAGCCAGGCGCTGCCCAT 245
   ::::: ||||| ||||| |||||
Db 56 AlaGlyThrThrTrpThrGlnGluIleValGluLeuGlnAsnGluGlyAspValGlu 75
   ::::: ||||| ||||| |||||
Qy 246 GAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGTCTCTGGAGTACCCACAGCCG--- 302
   ::::: ||||| ||||| |||||
Db 76 LysSerLysArgAlaProThrHisGlnArgPheProPheLeuGluMetLysIleProSer 95
   ::::: ||||| ||||| |||||
Qy 303 -----GGCTGGACATCATCAGAACTGACCTCCCGCTCATCAAGAGCCAC 353
   ::::: ||||| ||||| |||||
Db 96 LeuGlySerGlyLeuGluGlnAlaHisAlaMetProSerProArgIleLeuLysThrHis 115
   ::::: ||||| ||||| |||||
Qy 354 CTGCCCTACCGCTTCTCCCTCTGACCTCCACAAATGGAGACTCCAAAGGTTCATCTATATG 413
   ::::: ||||| ||||| |||||
Db 116 LeuProPheHisLeuLeuProSerLeuLeuGluLysAsnCysLysIleIleTyrVal 135
   ::::: ||||| ||||| |||||
Qy 414 GCTCGCAACCCCAAGGATCTGGTGTGCTTATTATCATGTTCCACCGCTCTCTCGGACC 473
   ::::: ||||| ||||| |||||
Db 136 AlaArgAsnProLysAspAsnMetValSerTyrHisPheGlnArgMetAsnLysAla 155
   ::::: ||||| ||||| |||||
Qy 474 ATGAGCTACCGAGCAGCCTTTCAAGAACTTCGCGGAGGTTTATGAATGATAAGCTGGC 533
   ::::: ||||| ||||| |||||
```





CC assays to identify small molecules that modulate or inhibit, for example, CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also CC used as in chromosome mapping, tissue typing, preventive medicine and CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein CC of the invention.

XX  
SQ Sequence 296 AA;

Alignment Scores:  
Pred. No.: 5.89e-33 Length: 296  
Score: 447.00 Matches: 98  
Percent Similarity: 54.4% Conservative: 57  
Best Local Similarity: 34.4% Mismatches: 108  
Query Match: 10.2% Indels: 22  
DB: 5 Gaps: 5

US-10-768-158-1 (1-2419) x ADI17132 (1-296)

```
QY 66 GAGACGAGTACTTCGAGTTCATCGGCTGGCGTGGCGCCCTTCTGCGCGGGAAGATG 125
D 10 GlnThrLysLeuLysGluValAlaGlyLeuProLeuArgAspSerThrValAspAsnTrp 29
QY 126 GAGGAGATCGCAACTTCCCGGTGGCGCCAGCAGCGTGGATCGTCACCTACCCCAAG 185
D 30 SerGlnLeuGlnThrPheLysAlaLysProAspLeuLeuLeuCysThrTyrProLys 49
QY 186 TCCGGCACCACTGCTGTCAGAGGTGGTCTACTTGGTGACCGGGCGCTGACCCCGAT 245
D 50 SerGlyThrTrpIleGlnGluValAlaAsnMetIleGluGlnAsnGlyAspValGlu 69
QY 246 GAGATCGGCTGTGAACATCGACGAGCATCCCGTCTCGGAGTAC-----CCA 296
D 70 LysCysGlnArgThrIleGlnHisArgHisProPheIleGluTrpAlaArgProPro 89
QY 297 CAGCG--GSCCTGGACATCATCAAGAACTGACCTCTCCCGCCTCATCAAGAGCAC 353
D 90 GlnProSerGlyValAspLysAlaAsnAlaMetProAlaProArgIleLeuArgHis 109
QY 354 CTGCCCTACCGCTTCTGCGCTCTGACCTCCACAAATGGAGACTCCAAAGTCACTATATG 413
D 110 LeuProIleGlnLeuLeuProProSerPheThrThrAsnAsnCysLysTyrLeuVal 129
QY 414 GCTCGAACCCAGGATCTGGTGTCTATTATCAGTTCCACCGCTCTCTCGGACC 473
D 130 AlaArgAsnAlaLysAspCysMetValSerPheTyrHisPheTyrArgMetCysGlnVal 149
QY 474 ATGAGCTACCGAGCAGCTTCAAGATTCGCGGAGGTTTATGAATGATAGCTGGC 533
D 150 LeuProAsnProGlyThrTrpAsnGluTyrPheGluThrPheIleAsnGlyLysValSer 169
QY 534 TACGGCTCTGTTGAGCAGCTGCAGAGTCTTGGGAGCACCGCATGGACTCGAAGCTG 593
D 170 TrpGlySerCysPheAspHisValLysGlyTrpTrpGluIleArgAspArgTyrGluIle 189
QY 594 CTTTTCTCAAGTATGACATCATCGGACCTGGTACGATGGTGAGGACGCTGGCC 653
D 190 LeuPheLeuPheTyrGluAspMetLysArgAspProLysArgGluIleGlnLysValMet 209
QY 654 AGATTCTGGGGGTGC-----TGTCACAGGCGCCAGCTGGAGCCCTG 698
D 210 GlnPheMetGlyLysAsnLeuAspGluValValAspLysIleValLeuGluThrSer 229
QY 699 ACGGAGCAGCTGCCACGAGCTGGTGACCGACGTCGCAACGCTGAGGCGCTGCC- 752
D 230 PheGluLysMetLysAsp-----AsnProLeuThrAsnPheSerThrIleProLysThr 247
QY 753 -----GTGGCGCGGGAAGAGTTGGCTGTGAAGGAC 785
D 248 IleMetAspGlnSerIleSerProPheMetArgLysGlyIleValGlyAspTrpLysAsn 267
QY 786 ATCTTCCCGTCTCCATGATGAGAGTTTGATGTTGTTGTTATTAACACAGAGATGGAAAG 845
D 268 HisPheThrValAlaGlnAsnGluArgPheAspGluIleTyrGluGlnLysMetAspGly 287
```

QY 846 TGTGACCTCACGTTT 860  
D 288 ThrSerLeuAsnPhe 292

## RESULT 48

ABM84002  
ID ABM84002 standard; protein; 327 AA.

XX AC ABM84002;

XX DT 18-NOV-2004 (first entry)

XX DE Human diagnostic and therapeutic pprotein SEQ ID NO:4251.

XX KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.

XX OS Homo sapiens.

XX PN WO2004023973-A2.

XX PD 25-MAR-2004.

XX PF 12-SEP-2003; 2003WO-US028227.

XX PR 12-SEP-2002; 2002US-0410259P.

XX PR 12-SEP-2002; 2002US-0410260P.

XX PA (INCY-) INCYTE CORP.

XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;  
PI Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;  
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;  
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;  
PI Peralta CH, Anderson SB, Rioux P, Shen BJ, Wu MC, Stuve LL;  
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;  
PI Xu Y, Kwong M, Policky JL, Hurwitz BU, Ma Y, Jackson JL, Gietzen D;  
PI Patury S, Shi X, Suarez CJ;

XX WPI; 2004-329368/30.

XX DR N-PSDB; ACN42654.

XX PT New diagnostic and therapeutic polynucleotides and polypeptides, useful  
PT in diagnosing a condition, disease or disorder associated with human  
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or  
PT in gene mapping.

XX PS Claim 27; Page: 190pp; English.

XX CC The invention relates to novel diagnostic and therapeutic polynucleotides  
CC selected from one of the 2722 sequences defined in the specification. A  
CC polynucleotide of the invention may have a use in gene therapy. The human  
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be  
CC used to diagnose a particular condition, disease or disorder associated  
CC with human molecules, e.g. cell proliferative disorders,  
CC autoimmune/inflammatory disorder, developmental disorder, endocrine  
CC disorder, neurological disorders, gastrointestinal disorders, or  
CC infections caused by virus, bacteria, fungi or parasite. The dithp  
CC molecules may also be used in genetic mapping, in identifying individuals  
CC from minute biological samples, in detecting single nucleotide  
CC polymorphisms, as molecular weight markers, and for somatic or germline  
CC gene therapy. The present sequence represents a dithp protein of the  
CC invention. Note: The sequence data for this patent is not represented in  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)

XX SQ Sequence 327 AA;

Alignment Scores:

Pred. No.: 6.2e-33 Length: 327  
Score: 447.00 Matches: 101  
Percent Similarity: 48.9% Conservative: 51  
Best Local Similarity: 32.5% Mismatches: 101





```
Db 77 ValArgValProPheLeuGluValAsnAspProGlyGluPro***GlyLeuGluThrLeu 96
      :::::|||||
Qy 318 AAGAACTGACCTCTCCCGCCTCATCAAGAGCACCTGCTACCGCTTTCCTCCCTCT 377
      |||:::
Db 97 LysAspThrProProArgLeuIleLysSerHisLeuProLeuAlaLeuLeuProGln 116
      |||:::
Qy 378 GACCTCCACATGAGACTCCCAAGCTCATCTATATGCTGCTGCAACCCACAGGATCGGTG 437
      |||:::
Db 117 ThrLeuLeuAspGlnLysValValValValValAlaArgAsnProLysAspValAla 136
      |||:::
Qy 438 GTGCTTATTATCAGTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGCACCTTCA 497
      |||:::
Db 137 ValSerTyrTyrHisPheHisArgMetClnLysAlaHisProGluProGlyThrTrpAsp 156
      |||:::
Qy 498 GAATTCGCCGAGGTTTATGAATGATAAGCTGGGC----- 533
      |||:::
Db 157 SerPheLeuGluLysPheMetAlaGlyGluGlyLeuAspTyrArgLysGluGlyVal 176
      |||:::
Qy 533 ----- 533
Db 177 LysProArgGlyGlyTyrAsnValGlnGlnProCysValGlyAlaProCysProLeu 196
      |||:::
Qy 534 -----TACGGCTCTCGTGTGAGCACGTGACAGAGTCTCTGGAGCACCGCATG 581
      |||:::
Db 197 Leu***ValSerTyrGlySerTyrTyrGlnHisValGlnGluTyrTrpGluLeuSerArg 216
      |||:::
Qy 582 GACTCGAACGCTTTTCTCAAGTATGAACATGATCGACCGGACCTGGTGACGATGGTG 641
      |||:::
Db 217 ThrHisProValLeuTyrLeuPheTyrGluAspMetLysGluAsnProLysArgGluIle 236
      |||:::
Qy 642 GAGCAGCTGGCCAGATCTCTGGGGTGTCTGTGACAGGCCAGCTGGAAGCCTGACG 701
      |||:::
Db 237 GlnLysIleLeuGluPheValGlyArgSerLeuProGluGluThrMetAspPheMetVal 256
      |||:::
Qy 702 GAGCAGTGC----- 710
Db 257 GlnHisThrSerPheLysGluMetLysLysAsnProMetThrAsnTyrThrThrValPro 276
      |||:::
Qy 711 CACGAGCTGTGGACCAAGTGTGCAACGCTGAGGCCCTGCCC---GTGGCCCGGGGAAGA 767
      |||:::
Db 277 GlnGluLeuMetAspHisSerIleSer-----ProPheMetArgLysGlyMet 292
      |||:::
Qy 768 GTTGGCTGTGGAGGACATCTTCCCGCTCTCCATGATGAGAAGTTTGACTTGGTGAT 827
      |||:::
Db 293 AlaGlyAspTyrLysThrThrPheThrValAlaGlnAsnGluArgPheAspAlaAspTyr 312
      |||:::
Qy 828 AAACAGAAGATGGGAAAGTGTGACCTCACGTTT 860
      |||:::
Db 313 AlaGluLysMetAlaGlyCysSerLeuSerPhe 323
      |||:::
RESULT 51
ABM84003
ID ABM84003 standard; protein; 327 AA.
XX
AC ABM84003;
XX
XX 18-NOV-2004 (first entry)
XX
XX Human diagnostic and therapeutic pprotein SEQ ID NO:4252.
XX
XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
XX Homo sapiens.
XX
XX WO2004023973-A2.
XX
XX 25-MAR-2004.
XX
XX 12-SEP-2003; 2003WO-US028227.
XX
XX 12-SEP-2002; 2002US-0410259P.
XX
XX 12-SEP-2002; 2002US-0410260P.
XX
```

```
XX
PA (INCY-) INCYTE CORP.
XX
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
XX Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
XX Mooney EM, Deleage AM, Panesar IS, Banville SC, Reddy TP;
XX Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
XX Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
XX Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtan ES;
XX Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
XX Patry S, Shi X, Suarez CJ;
XX
XX WPI: 2004-329368/30.
XX N-PSDB: ACN42655.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
XX in diagnosing a condition, disease or disorder associated with human
XX molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX in gene mapping.
XX
XX Claim 27; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
XX selected from one of the 2722 sequences defined in the specification. A
XX polynucleotide of the invention may have a use in gene therapy. The human
XX diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
XX used to diagnose a particular condition, disease or disorder associated
XX with human molecules, e.g. cell proliferative disorders,
XX autoimmune/inflammatory disorders, gastrointestinal disorders, or
XX infections caused by virus, bacteria, fungi or parasite. The dithp
XX molecules may also be used in genetic mapping, in identifying individuals
XX from minute biological samples, in detecting single nucleotide
XX polymorphisms, as molecular weight markers, and for somatic or germline
XX gene therapy. The present sequence represents a dithp protein of the
XX invention. Note: The sequence data for this patent is not represented in
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
XX SQ Sequence 327 AA;
XX
XX Alignment Scores:
XX Pred. No.: 6.2e-33 Length: 327
XX Score: 447.00 Matches: 101
XX Percent Similarity: 48.9% Conservative: 51
XX Best Local Similarity: 32.5% Mismatches: 101
XX Query Match: 10.2% Indels: 58
XX DB: 8 Gaps: 5
XX
XX US-10-768-158-1 (1-2419) x ABM84003 (1-327)
Qy 90 GCGTGTGGTGTGCGCGCTTCTGCGCGGGAAGATGGAGAGATCGGCACTTCCCGTG 149
      |||:::
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
      |||:::
Qy 150 CGGCCAGCGACGTGTGGATCGTCACTACCCAGTCCCGCACCGCTGCTGCAGAG 209
      |||:::
Db 37 ArgProAspAspLeuLeuIleAsnThrTyrProLysSerGlyThrThrTrpValSerGln 56
      |||:::
Qy 210 GTGCTTACTTGTGTGAGCCAGGCGCTGACCCCATGATCGGCTTGTGAACATCGAC 269
      |||:::
Db 57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysAsnArgAlaProIleTyr 76
      |||:::
Qy 270 GAGCAGCTCCCGCTCTCCCGCTCATCAAGAGCACCTGCGCTACCGCTTTCGCCCTCT 317
      |||:::
Db 77 ValArgValProPheLeuGluValAsnAspProGlyGluProSerGlyLeuGluThrLeu 96
      |||:::
Qy 318 AAGAACTGACCTCTCCCGCCTCATCAAGAGCACCTGCGCTACCGCTTTCGCCCTCT 377
      |||:::
Db 97 LysAspThrProProArgLeuIleLysSerHisLeuProLeuAlaLeuLeuProGln 116
      |||:::
Qy 378 GACCTCCACATGAGACTCCCAAGCTCATCTATATGCTGCTGCAACCCACAGGATCGGTG 437
      |||:::
XX
```

Db 117 ThrLeuLeuAapGlnLysValValValValValValValValValValValValValValValVal 136  
Qy 438 GTGCTTATTATCAGTCCACCGCTCTCTGCGGACCATGAGTACCGAGGACACCTTTCAA 497  
Db 137 ValSerTyrTyrHisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAsp 156  
Qy 498 GAATTCCTCCGAGGTTTATGAATGATTAAGCTGGGC----- 533  
Db 157 SerPheLeuGluLysPheMetAlaGlyGlyLeuAaspTApArgLysGluGlyVal 176  
Qy 533 ----- 533  
Db 177 LysProArgGlyGlyTyrAsnValGlnGlnProCysValGlyAlaProCysProLeu 196  
Qy 534 -----TACGGCTCTGTTTGGACACGTGACAGGAGTTCTGGGACGACCGCATG 581  
Db 197 Leu\*\*ValSerTyrGlySerTrpTyrGlnHisValGlnGluTrpTrpGluLeuSerArg 216  
Qy 582 GACTCGAACGTCTTTTCTCAAGTATGAACATGATCGGACCTGGTGACGATGGTG 641  
Db 217 ThrHisProValLeuTyrLeuPheTyrGluAspMetLysGluAsnProLysArgGluIle 236  
Qy 642 GAGCAGCTGGCCAGATCTCTGGGGTGTCTGTGACAGGCCAGCTGGAAGCCCTGACG 701  
Db 237 GlnLysIleLeuGluPheValGlyArgSerLeuProGluGluThrMetAspPheMetVal 256  
Qy 702 GAGCACTGC----- 710  
Db 257 GlnHisThrSerPheLysGluMetLysLysAsnProMetThrAsnTyrThrValPro 276  
Qy 711 CACGAGCTGTGGACGAGTGTGCAACGCTGAGGCCCTGCCC---GTGGGCGCGGGAAGA 767  
Db 277 GlnGluLeuMetAspHisSerIleSer-----ProPheMetArgLysGlyMet 292  
Qy 768 GTTGGGCTGTGGAGACATCTTCACCGTCTCCATGATGAGAGTTTGACTTGGTGTAT 827  
Db 293 AlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPheAspAlaAspTyr 312  
Qy 828 AACAGAGATGGGAAAGTGTGACCTCACGTTT 860  
Db 313 AlaGluLysMetAlaGlyCysSerLeuSerPhe 323

RESULT 52  
ABM84004  
ID ABM84004 standard; protein; 327 AA.  
XX  
AC ABM84004;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human diagnostic and therapeutic pprotein SEQ ID NO:4253.  
XX  
XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.  
XX  
XX Homo sapiens.  
XX  
XX WO2004023973-A2.  
XX  
XX 25-MAR-2004.  
XX  
XX 12-SEP-2003; 2003MO-US028227.  
XX  
XX 12-SEP-2002; 2002US-0410259P.  
XX  
XX 12-SEP-2002; 2002US-0410260P.  
XX  
XX (INCY-) INCYTE CORP.  
XX  
XX Schmidt JP, Wright RJ, Bruns CM, Marianovic MM, Shen F;  
XX Harthshorne TA, Suchorski MT, Altus CM, Pitts SJ, Elder LV;  
XX Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;  
XX Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH;  
XX Peralca CH, Anderson SB, Rioux P, Shen EU, Wu MC, Stuve LU;  
XX Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtton ES;

PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;  
PI Patury S, Shi X, Suarez CJ;  
XX WPI: 2004-329368/30.  
DR N-PSDB, ACN42656.  
XX  
PT New diagnostic and therapeutic polynucleotides and polypeptides, useful  
PT in diagnosing a condition, disease or disorder associated with human  
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or  
PT in gene mapping.  
XX  
PS Claim 27; Page: 190pp; English.  
XX  
CC The invention relates to novel diagnostic and therapeutic polynucleotides  
CC selected from one of the 2722 sequences defined in the specification. A  
CC polynucleotide of the invention may have a use in gene therapy. The human  
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be  
CC used to diagnose a particular condition, disease or disorder associated  
CC with human molecules, e.g. cell proliferative disorders,  
CC autoimmune/inflammatory disorder, developmental disorders, endocrine  
CC disorder, neurological disorders, gastrointestinal disorders, or  
CC infections caused by virus, bacteria, fungi or parasite. The dithp  
CC molecules may also be used in genetic mapping, in identifying individuals  
CC from minute biological samples, in detecting single nucleotide  
CC polymorphisms, as molecular weight markers, and for somatic or germline  
CC gene therapy. The present sequence data for this patent is not represented in  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm  
XX  
SQ Sequence 327 AA;  
Alignment Scores:  
Pred. No.: 6.2e-33 Length: 327  
Score: 447.00 Matches: 101  
Percent Similarity: 48.9% Conservative: 51  
Best Local Similarity: 32.5% Mismatches: 101  
Query Match: 10.2% Indels: 58  
DB: Gaps: 5  
US-10-768-158-1 (1-2419) x ABM84004 (1-327)  
Qy 90 GGGCTGGCTGCGCGCTTCTGCGCGGGAAGATGAGGAGATCGCAACTTCCCGGTG 149  
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36  
Qy 150 CGGCCACGACGCTGTGGATCGTACCTACCCCAAGTCCGCGACACGCTTGTCTGAGGAG 209  
Db 37 ArgProAspAspLeuLeuIleAsnThrTyrProLysSerGlyThrThrTrpValSerGln 56  
Qy 210 GTGCTCTACTTGTGAGCGGCGCTGACCCGATGAGATCGGCTTGTGAACTCGAC 269  
Db 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysAsnArgAlaProIleTyr 76  
Qy 270 GAGCAGCTCCCGTCTCGGAGTACCCACACCGC-----GGCTGGACATCATC 317  
Db 77 ValArgValProPheLeuGluValAsnAspProGlyGluProSerGlyLeuGluThrLeu 96  
Qy 318 AAGGAAGTACCTCTCCCGCTCATCAAGAGCCACCTGCGCTTCTGCTCTCT 377  
Db 97 LysAspThrProProArgLeuIleLysSerHisLeuProLeuAlaLeuLeuProGln 116  
Qy 378 GACCTCCACAATGAGACTCAAGGTCTATATGCTGCGCAACCCCAAGGATCTGGT 437  
Db 117 ThrLeuLeuAspGlnLysValValValValValValValValValValValValVal 136  
Qy 438 GTGCTTATTATCAGTTCACCGCTCTCTGCGGACCATGAGTACCGAGGACACCTTTCAA 497  
Db 137 ValSerTyrTyrHisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAsp 156  
Qy 498 GAATTCCTCCGAGGTTTATGAATGATTAAGCTGGGC----- 533  
Db 157 SerPheLeuGluLysPheMetAlaGlyGlyLeuAaspTApArgLysGluGlyVal 176

QY 533 ----- 533

Db 177 LysProArgGlyGlyTyrAsnValGlnGlnProCysValGlyAlaProCysProLeu 196

QY 534 -----TACGGCTCTGTTGAGCACGTGAGGAGTTCCTGGAGCACCAGCATG 581

Db 197 Leu\*\*\*ValSerTyrGlySerTrpTyrGlnHisValGlnGluTrpTpgGluLeuSerArg 216

QY 582 GACTCGAACGTCTTTTCTCAAGATGAAAGACATGATCGGACCTGGTGACATGGTG 641

Db 217 ThrHisProValLeuTyrLeuPheTyrGluAspMetLysGluAsnProLysArgGluIle 236

QY 642 GAGCAGCTGGCCATCTCTGGGGTGTCTGTGACAAAGCCAGCTGGAAGCCCTGACG 701

Db 237 GlnLysIleLeuGluPheValGlyArgSerLeuProGluGluThrMetAspPheMetVal 256

QY 702 GAGCACTGCG----- 710

Db 257 GlnHisThrSerPheLysGluMetLysLysAsnProMetThrAsnTyrThrValPro 276

QY 711 CACGAGCTGGTGGACCGATGTCGAACGCTGAGGCCCTGCCC---GTGGGCCGGGGAAGA 767

Db 277 GlnGluLeuMetAspHisSerIleSer-----ProPheMetArgLysGlyMet 292

QY 768 GTTGGGCTGTGGAAGGACATCTTCACCGTCTCCATGAATGAGAAAGTTTGACTTGGTGAT 827

Db 293 AlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPheAspAlaAspTyr 312

QY 828 AAACAGAAGTGGGAAAGTGTGACCTCACGCTT 860

Db 313 AlaGluLysMetAlaGlyCysSerLeuSerPhe 323

RESULT 53

AD117319

ID AD117319 standard; protein; 269 AA.

XX AC AD117319;

XX DT 15-APR-2004 (first entry)

XX DE Polypeptide homologous to a human NOVX domain SeqID 855.

XX KW NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes; inflammation;

XX KW autoimmune disorder; allergy; blood disorder;

XX KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;

XX KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;

XX KW Alzheimer's disease; infection; str.

XX OS Unidentified.

XX OS WO200268649-A2.

XX PN 06-SEP-2002.

XX PD 31-JAN-2002; 2002WO-US002785.

XX PF 31-JAN-2001; 2001US-0265395P.

XX PR 31-JAN-2001; 2001US-0265412P.

XX PR 31-JAN-2001; 2001US-0265514P.

XX PR 31-JAN-2001; 2001US-0265517P.

XX PR 02-FEB-2001; 2001US-0266406P.

XX PR 05-FEB-2001; 2001US-0266767P.

XX PR 07-FEB-2001; 2001US-0266975P.

XX PR 07-FEB-2001; 2001US-0267057P.

XX PR 08-FEB-2001; 2001US-0267459P.

XX PR 09-FEB-2001; 2001US-0267823P.

XX PR 15-FEB-2001; 2001US-0268974P.

XX PR 26-FEB-2001; 2001US-0271664P.

XX PR 27-FEB-2001; 2001US-0271839P.

XX PR 27-FEB-2001; 2001US-0271855P.

XX PR 02-MAR-2001; 2001US-0272788P.

XX PR 02-MAR-2001; 2001US-0273046P.

PR 14-MAR-2001; 2001US-0275925P.

PR 14-MAR-2001; 2001US-0275947P.

PR 14-MAR-2001; 2001US-0275950P.

PR 14-MAR-2001; 2001US-0275989P.

PR 15-MAR-2001; 2001US-0276448P.

PR 15-MAR-2001; 2001US-0276450P.

PR 16-MAR-2001; 2001US-0276397P.

PR 16-MAR-2001; 2001US-0276768P.

PR 20-MAR-2001; 2001US-0278652P.

PR 26-MAR-2001; 2001US-0278775P.

PR 26-MAR-2001; 2001US-0278778P.

PR 29-MAR-2001; 2001US-0279882P.

PR 29-MAR-2001; 2001US-0279884P.

PR 30-MAR-2001; 2001US-0280147P.

PR 11-APR-2001; 2001US-0282992P.

PR 11-APR-2001; 2001US-0283083P.

PR 20-APR-2001; 2001US-0285133P.

PR 23-APR-2001; 2001US-0285743P.

PR 03-MAY-2001; 2001US-0288327P.

PR 03-MAY-2001; 2001US-0288504P.

PR 29-MAY-2001; 2001US-0294047P.

PR 30-MAY-2001; 2001US-0294473P.

PR 08-JUN-2001; 2001US-0296964P.

PR 18-JUN-2001; 2001US-0298959P.

PR 19-JUN-2001; 2001US-0299324P.

PR 13-AUG-2001; 2001US-0312020P.

PR 16-AUG-2001; 2001US-0312889P.

PR 16-AUG-2001; 2001US-0312908P.

PR 21-AUG-2001; 2001US-0313390P.

PR 28-AUG-2001; 2001US-0315470P.

PR 31-AUG-2001; 2001US-0316447P.

PR 07-SEP-2001; 2001US-0318115P.

PR 07-SEP-2001; 2001US-0318118P.

PR 12-SEP-2001; 2001US-0318740P.

PR 19-SEP-2001; 2001US-0323379P.

PR 18-OCT-2001; 2001US-0330245P.

PR 18-OCT-2001; 2001US-0330308P.

PR 14-NOV-2001; 2001US-0332701P.

XX XX

PA (CURA-) CURAGEN CORP.

XX PI Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shinkets RA;

XX PI Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;

PI Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CEA;

PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;

XX WPI; 2002-706998/76.

DR New NOVX polypeptides and nucleic acids, useful for preventing or

XX treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,

XX atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or

XX pharmacogenomics.

PS Disclosure; SEQ ID NO 855; 1498pp; English.

XX XX

CC This invention relates to a novel nucleic acids, and encoded polypeptides

CC thereof, which have properties related to the stimulation of biochemical

CC or physiological responses in a cell, tissue, organ or organism.

CC Specifically, it refers to the use of biologically active fragments for

CC diagnostic and prognostic assays and furthermore in the treatment of

CC diverse pathological conditions. The present invention describes novel

CC human and murine NOVX proteins, as well as methods to modulate their

CC expression using antisense oligos, ribozymes and peptide nucleic acids.

CC The NOVX polypeptides, polynucleotides and antibodies are useful in

CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,

CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in

CC treating or preventing diseases such as inflammation, autoimmune

CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome

CC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,

CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy

CC and epilepsy. Accordingly, these molecules have many activities including

CC cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,

CC haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,

CC

CC antiasthmatic, nephrotropic, antiarthritic, hepatotropic,  
CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,  
CC relaxant and anticonvulsant. In addition, they are useful in screening  
CC assays to identify small molecules that modulate or inhibit, for example,  
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also  
CC used as in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. This polypeptide is a protein fragment that is  
CC homologous to a domain of a human NOVX protein of the invention.  
XX  
SQ Sequence 269 AA;

Alignment Scores:  
Pred. No.: 1.08e-32 Length: 269  
Score: 444.00 Matches: 104  
Percent Similarity: 55.9% Conservative: 39  
Best Local Similarity: 40.6% Mismatches: 83  
Query Match: 10.1% Indels: 30  
DB: 5 Gaps: 7

US-10-768-158-1 (1-2419) x ADI17319 (1-269)

Qy 141 TTCCCGTGCAGCCAGCGAGTGTGGATCGTCACTACCCAGTCCGGACCAAGCTTG 200  
Db PheGlnAlaArgProAspValLeuLeuAlaGlyTyrProLysSerGlyThrTrp 39  
Qy 201 CTGAGGAGGTGTCTACTTGTGTAGCCAGGCGCTGAC-----CCCGAT 245  
Db LeuGlnGluLeuLeuHisProAsnValGlyAspPheGluProSerProSerAsp 59  
Qy 246 GAGATCGCTGTAGTAACATCGACGAGCAGCTCCGGTCTGGAGTACCCACGCCGGC 305  
Db ProLeuLeuPheArgAsn-----ProTrpLeuGluTyrProLys---Gly 73  
Qy 306 CTGACATCATCAAGAACTG-----ACCTCTCCCGCTCATCAAGAGCCAC 353  
Db GluAspTrpTyrGluThrLeuLysProMetProSerSerProArgLeuLeuLysThrHis 93  
Qy 354 CTGCCCTACCGCTTCTCCCTCTGACCTCCACATGGAGACTCCAGTTCATCATATG 413  
Db LeuProLeuGluLeuLeuProLysSerPheLeuSerSerLysAlaLysIleIleTyrVal 113  
Qy 414 GCTCGACACCCAGGACCTTTCAAGAAATTCGCCGAGGTTTATGAATGATAAGCTG 473  
Db LeuProAlaAspProGlyThrPheGluGluPheLeuGluAlaPheLeuAsnGlyLysVal 153  
Qy 531 GGCTACGCTCCTGGTTGACACGTCGAGAGTTCGGGAGCACCCGATCGACTCGAAC 590  
Db LeuTyrGlySerTyrPheAspHisValLeuGlyTyrTrpGluLeuArgProGluProGln 173  
Qy 591 GTGCTTTTTCAGATATGAACATGATCGGACCTGGGACCTGGTGCAGTGTGGAGCAGCTG 650  
Db ValLeuPheLeuAspTyrGluAspLysGluAspProAlaGlyGluLeuLysIle 193  
Qy 651 GCCAGATCTCGGGGTCTCTGTGACAAAGCCAGCTGGAAGCCCTCAGCAGAGCAGTGC 710  
Db AlaGluPheLeuGlyLeuProLeuSerGluGluGluLeuAspLysLeuLeuAspHisSer 213  
Qy 711 CAC-----CAGCTGGTGGACAGTGTGCAACGCTGAGGCCCTCGCCGTGGGC--- 758  
Db SerPhePheLeuMetLysLeuAsnProLeuSerAsnTyrGluThrLeuCysLeuGlyLys 233  
Qy 759 -----CGGGAAGAGTTGGCTCTGGAAGACATC 788  
Db SerLysGlyArgLysSerProPheMetArgLysGlyLeuValGlyAspTrpLysAsnTyr 253  
Qy 789 TTCACCGTCTCCATGAATGAGAAGTTTCATTTGGTGTATTAACAGAAAG 836  
Db PhetrProGluGlnAsnGluLysPheAspLysValIleLysGluLys 269

RESULT 54  
ADI17331  
ID ADI17331 standard; protein; 269 AA.  
XX  
AC ADI17331;  
XX  
DT 15-APR-2004 (first entry)  
XX  
DE Polypeptide homologous to a human NOVX domain SeqID 867.  
XX  
KW NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes; inflammation;  
KW autoimmune disorder; allergy; blood disorder;  
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;  
KW immunoglobulin (IgA nephropathy; cirrhosis; arthritis;  
XX Alzheimer's disease; infection; str.  
OS Unidentified.  
XX  
PN W0200268649-A2.  
XX  
PD 06-SEP-2002.  
XX  
XX 31-JAN-2002; 2002WO-US002785.  
PR 31-JAN-2001; 2001US-0265395P.  
PR 31-JAN-2001; 2001US-0265412P.  
PR 31-JAN-2001; 2001US-0265514P.  
PR 31-JAN-2001; 2001US-0265517P.  
PR 02-FEB-2001; 2001US-0266406P.  
PR 05-FEB-2001; 2001US-0266787P.  
PR 07-FEB-2001; 2001US-0266975P.  
PR 07-FEB-2001; 2001US-0267057P.  
PR 08-FEB-2001; 2001US-0267459P.  
PR 09-FEB-2001; 2001US-0267823P.  
PR 15-FEB-2001; 2001US-0268974P.  
PR 26-FEB-2001; 2001US-0271664P.  
PR 27-FEB-2001; 2001US-0271839P.  
PR 02-MAR-2001; 2001US-0272788P.  
PR 02-MAR-2001; 2001US-0273046P.  
PR 14-MAR-2001; 2001US-0275925P.  
PR 14-MAR-2001; 2001US-0275947P.  
PR 14-MAR-2001; 2001US-0275950P.  
PR 14-MAR-2001; 2001US-0275989P.  
PR 15-MAR-2001; 2001US-0276448P.  
PR 15-MAR-2001; 2001US-0276450P.  
PR 16-MAR-2001; 2001US-0276397P.  
PR 16-MAR-2001; 2001US-0276768P.  
PR 20-MAR-2001; 2001US-0278652P.  
PR 26-MAR-2001; 2001US-0278775P.  
PR 26-MAR-2001; 2001US-0278778P.  
PR 29-MAR-2001; 2001US-0279882P.  
PR 29-MAR-2001; 2001US-0279884P.  
PR 30-MAR-2001; 2001US-0280147P.  
PR 11-APR-2001; 2001US-0282992P.  
PR 11-APR-2001; 2001US-0283083P.  
PR 20-APR-2001; 2001US-0285133P.  
PR 23-APR-2001; 2001US-0285749P.  
PR 03-MAY-2001; 2001US-0288327P.  
PR 03-MAY-2001; 2001US-0288504P.  
PR 30-MAY-2001; 2001US-0294473P.  
PR 30-MAY-2001; 2001US-0294047P.  
PR 18-JUN-2001; 2001US-0298959P.  
PR 19-JUN-2001; 2001US-0299324P.  
PR 13-AUG-2001; 2001US-0312020P.  
PR 16-AUG-2001; 2001US-0312889P.  
PR 16-AUG-2001; 2001US-0312908P.  
PR 21-AUG-2001; 2001US-0313390P.  
PR 28-AUG-2001; 2001US-0315470P.  
PR 31-AUG-2001; 2001US-0316447P.  
PR 07-SEP-2001; 2001US-0318115P.  
PR 07-SEP-2001; 2001US-0318118P.











```

Db 186 ArgValLeuPheLeuPheTyrGluAspLeuLysGluAspIleArgLysGluValIleLys 205
Qy 648 CTGCCAGATTCTCTGGGGTGTCTCTGTACAAAGCCCGAGCTGGGAAGCCCTGACCGAGCAC 707
Db 206 LeuIleHisPheLeuGluArgLysProSerGluGluLeuValAspArgIleHisHis 225
Qy 708 TGCAC-----CAGCTGGTGGACGAC 728
Db 226 ThrSerPheGlnGluMetLysAsnProSerThrAsnTyrThrLeuProAspGlu 245
Qy 729 TGCTGCAACGCTGAGCCCTCC---GTGGCCGGGGAAGAGTTGGGCTGTGGAAAGAC 785
Db 246 IleMetLanGlnLysLeuSerProPheMetArgLysGlyIleThrGlyAspTrpLysAsn 265
Qy 786 ATCTTCACCGTCTCCATGAATGAGAAGTTTCACTTGGTGTATAACAGAAAGATGGGAAG 845
Db 266 HisPheThrValAlaLeuAsnGluLysPheAspLysHisTyrGluGlnMetLysGlu 285
Qy 846 TGTGACCTCACGTTT 860
Db 286 SerThrLeuLysPhe 290

RESULT 59
ADD48896
ID ADD48896 standard; protein; 294 AA.
XX AC ADD48896;
XX XX
XX 02-DEC-2004 (revised)
XX DT 29-JAN-2004 (first entry)
XX XX
XX Human Protein P49888, SEQ ID NO 14607.
XX DE
XX KW Human; pain; neuronal tissue; gene therapy;
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;
XX KW spared nerve injury; SNI; Chung.
XX OS Homo sapiens.
XX OS Unidentified.
XX XX
XX PN WO2003016475-A2.
XX XX
XX PD 27-FEB-2003.
XX XX
XX PF 14-AUG-2002; 2002WO-US025765.
XX XX
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX XX
XX PA (GENO ) GEN HOSPITAL CORP.
XX PA (FARB ) BAYER AG.
XX XX
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX XX
XX DR WPI; 2003-268312/26.
XX DR GENBANK; P49888.
XX XX
XX New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX XX
XX PS Example 1; Page; 1017pp; English.
XX XX
XX The invention discloses a composition comprising two or more isolated rat
XX CC or human polynucleotides or a polynucleotide which represents a fragment,
XX CC derivative or allelic variation of the nucleic acid sequence. Also
XX CC claimed are a vector comprising the novel polynucleotide, a host cell
XX CC comprising the vector, a method for identifying a nucleotide sequence
XX CC which is differentially regulated in an animal subjected to pain and a
XX CC kit to perform the method, an array, a method for identifying an agent
XX CC that increases or decreases the expression of the polynucleotide sequence
XX CC that is differentially expressed in neuronal tissue of a first animal

```

```

CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (described in Table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 294 AA;

```

```

Alignment Scores:
Pred. NO.: 1.02e-31 Length: 294
Score: 434.00 Matches: 95
Percent Similarity: 53.0% Conservative: 56
Best Local Similarity: 33.3% Mismatches: 116
Query Match: 9.9% Indels: 18
DB: 7 Gaps: 3

```

US-10-768-158-1 (1-2419) x ADD48896 (1-294)

```

Qy 60 GAGTTCGAGAGCAAGTACTTCGAGTTCATGCGGTGGCGGCTTCGCGCCCTTCGCGCGGG 119
Db 6 AspTyrTyrGluLysPheGluGluValHisGlyIleLeuMetTyrLysAspPheValLys 25
Qy 120 AAGATGAGGAGATCGCCCAACTTCCCGGTGGCGGCCAGCAGCGTGTGGATCGTCACCTAC 179
Db 26 TyrTrpAsnValGluAlaPheGlnAlaArgProAspAspLeuValIleAlaThrTyr 45
Qy 180 CCAAGTCGCGCACCAGCTTGTCTCAGAGGTGTCTACTTGTGTGACCGCGGCGCTGAC 239
Db 46 ProlysSerGlyThrThrTrpValSerGluIleValTyrMetIleTyrLysGluGlyAsp 65
Qy 240 CCGATGAGATCGCGCTTGTGATGAACATCGACGAGCGTCCCGCTCTGGAGTACCCACAG 299
Db 66 ValGluLysCysLysLysGluAspValIlePheAsnArgIleProPheLeuGluCysArgLys 85
Qy 300 CCG-----GGCCTGGACATCATCAAGGAACCTGACCTCTCCCGGCTCATCAAG 347
Db 86 GluAsnLeuMetAsnGlyValLysGlnLeuAspGluMetAsnSerProArgIleValLys 105
Qy 348 AGCCACCTGCCCTACCGCTTCTGCCCTCTGACCTCCACATGGAGACTCCAGGTATC 407
Db 106 ThrHisLeuProGluLeuLeuProAlaSerPheTrpGluLysAspCysLysIleIle 125
Qy 408 TATATGGCTCGCAACCCCAAGGATCTGTGTGTCTTATTATCAGTTCCACCGCTCTCTG 467
Db 126 TyrLeuCysArgAsnAlaLysAspValAlaValSerPheTyrTyrPhePheLeuMetVal 145
Qy 468 CGGACCATGAGTACCGGAGCACCTTTCAAGAAATTTCTCCGAGGTTTATGAATGAAG 527
Db 146 AlaGlyHisProAsnProGlySerPheProGluPheValGluLysPheMetGlnGlyGln 165
Qy 528 CTGGGCTACGGCTCTGTTGACACGTCGAGGAGTCTGGGAGCACCGCATGGACTCG 587
Db 166 ValProTyrGlySerTrpTyrLysHisValLysSerTrpTrpGluLysGlyLysSerPro 185
Qy 588 AACGTGCTTTTCTCAGTATGAAGACATCATCGGACCTCGGTGACCGATGGTGAGCAG 647
Db 186 ArgValLeuPheLeuPheTyrGluAspLeuLysGluAspIleArgLysGluValIleLys 205
Qy 648 CTGGCCAGATTCTCTGGGGTGTCTCTGTGACAAGGCCCGCCAGGCTGAGCGGAGCAC 707

```

Db 206 LeuIleHisPheLeuGluArgLysProSerGluGluLeuValAspArgIleIleHisHis 225  
Qy 708 TGCCAC-----CAGCTGGTGGACGAC 728  
Db 226 ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrThrLeuProAspGlu 245  
Qy 729 TGCTGCACGCTGAGCCCTGCCCC---GTGGGCGGGGAAGAGTTGGCTGTGGAAGAC 785  
Db 246 IleMetAsnGlnLysLeuSerProPheMetArgLysGlyIleThrGlyAspTyrLysAsn 265  
Qy 786 ATCTTCACCGCTCCATGACAGTTTGACTTGGTGTTATTAACACAGACATGGGAAG 845  
Db 266 HisPheThrValAlaLeuAsnGluLysPheAspLysHisTyrGluGlnGlnMetLysGlu 285  
Qy 846 TGTGACCTCACGTTT 860  
Db 286 SerThrLeuLysPhe 290  
RESULT 60  
AD117049  
ID AD117049 standard; protein; 307 AA.  
XX  
AC AD117049;  
DT 15-APR-2004 (first entry)  
XX  
DE Chicken NOVX protein homologue SeqID 585.  
XX  
KW chicken; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;  
KW inflammation; autoimmune disorder; allergy; blood disorder;  
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;  
KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;  
KW Alzheimer's disease; infection; s.  
XX  
OS Gallus gallus.  
XX  
PN W0200268649-A2.  
XX  
PD 06-SEP-2002.  
XX  
PF 31-JAN-2002; 2002WO-US002785.  
XX  
PR 31-JAN-2001; 2001US-0265395P.  
PR 31-JAN-2001; 2001US-0265412P.  
PR 31-JAN-2001; 2001US-0265514P.  
PR 31-JAN-2001; 2001US-0265517P.  
PR 02-FEB-2001; 2001US-0266406P.  
PR 05-FEB-2001; 2001US-0266767P.  
PR 07-FEB-2001; 2001US-0266975P.  
PR 07-FEB-2001; 2001US-0267057P.  
PR 08-FEB-2001; 2001US-0267459P.  
PR 09-FEB-2001; 2001US-0267823P.  
PR 15-FEB-2001; 2001US-0268974P.  
PR 26-FEB-2001; 2001US-0271664P.  
PR 27-FEB-2001; 2001US-0271839P.  
PR 27-FEB-2001; 2001US-0271855P.  
PR 02-MAR-2001; 2001US-0272788P.  
PR 02-MAR-2001; 2001US-0273046P.  
PR 14-MAR-2001; 2001US-0275925P.  
PR 14-MAR-2001; 2001US-0275947P.  
PR 14-MAR-2001; 2001US-0275950P.  
PR 15-MAR-2001; 2001US-0275989P.  
PR 15-MAR-2001; 2001US-0276448P.  
PR 15-MAR-2001; 2001US-0276450P.  
PR 15-MAR-2001; 2001US-0276397P.  
PR 16-MAR-2001; 2001US-0276768P.  
PR 20-MAR-2001; 2001US-0278652P.  
PR 26-MAR-2001; 2001US-0278775P.  
PR 26-MAR-2001; 2001US-0278778P.  
PR 29-MAR-2001; 2001US-0279882P.  
PR 29-MAR-2001; 2001US-0279884P.  
PR 30-MAR-2001; 2001US-0280147P.  
PR 11-APR-2001; 2001US-0282992P.

PR 11-APR-2001; 2001US-0283083P.  
PR 20-APR-2001; 2001US-0285133P.  
PR 23-APR-2001; 2001US-0285749P.  
PR 03-MAY-2001; 2001US-0288327P.  
PR 03-MAY-2001; 2001US-0288504P.  
PR 29-MAY-2001; 2001US-0294047P.  
PR 30-MAY-2001; 2001US-0294473P.  
PR 08-JUN-2001; 2001US-0296964P.  
PR 18-JUN-2001; 2001US-0298959P.  
PR 19-JUN-2001; 2001US-0299324P.  
PR 13-AUG-2001; 2001US-0312020P.  
PR 16-AUG-2001; 2001US-0312889P.  
PR 21-AUG-2001; 2001US-0312908P.  
PR 28-AUG-2001; 2001US-0313390P.  
PR 31-AUG-2001; 2001US-0316447P.  
PR 07-SEP-2001; 2001US-0318115P.  
PR 07-SEP-2001; 2001US-0318118P.  
PR 12-SEP-2001; 2001US-0318740P.  
PR 19-SEP-2001; 2001US-0323379P.  
PR 18-OCT-2001; 2001US-0330245P.  
PR 18-OCT-2001; 2001US-0330308P.  
PR 14-NOV-2001; 2001US-0332701P.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
XX Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shinkets RA;  
PI Li L, Gangolli EA, Padigar M, Anderson DW, Rastelli L, Miller CE;  
PI Garlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CEA;  
PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;  
XX  
DR WPI; 2002-706998/76.  
XX  
XX New NOVX polypeptides and nucleic acids, useful for preventing or  
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy, typing or  
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX  
PS Disclosure; SEQ ID NO 585; 1498pp; English.  
XX  
XX This invention relates to a novel nucleic acids, and encoded polypeptides  
CC thereof, which have properties related to the stimulation of biochemical  
CC or physiological responses in a cell, tissue, organ or organism.  
CC Specifically, it refers to the use of biologically active fragments for  
CC diagnostic and prognostic assays and furthermore in the treatment of  
CC diverse pathological conditions. The present invention describes novel  
CC human and murine NOVX proteins, as well as methods to modulate their  
CC expression using antisense oligos, ribozymes and peptide nucleic acids.  
CC The NOVX polypeptides, polynucleotides and antibodies are useful in  
CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,  
CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in  
CC treating or preventing diseases such as inflammation, autoimmune  
CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome  
CC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,  
CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy  
CC and epilepsy. Accordingly, these molecules have many activities including  
CC cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,  
CC haemostatic, anti-HIV, antidiabetic, antiatherosclerotic, anorectic,  
CC antiasthmatic, nephrotropic, antiarthritic, hepatotropic,  
CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,  
CC relaxant and anticonvulsant. In addition, they are useful in screening  
CC assays to identify small molecules that modulate or inhibit, for example,  
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also  
CC used as in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein  
CC of the invention.  
XX  
SQ Sequence 307 AA;  
Alignment Scores:  
Pred. No.: 3,89e-31 Length: 307  
Score: 428.00 Matches: 87  
Percent Similarity: 55.4% Conservative: 61



```
DB: 4 Gaps: 5
US-10-768-158-1 (1-2419) x ABE12209 (1-304)
Qy 21 ATGCGGAGAGCGAGCGGCGAGACCCAGACCCCGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLeuAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy 78 TTCAGTTCATGGGTGGCGTGGCGCCCTTCTGCGCGGGAAGATGGAGGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuLeuSerLysGluTrpTrpGluLysValCys 40
Qy 138 AACTTCCCGTGGCGCGAGCGTGGATCGTACCTACCCAGATCGGCGACCGAGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTrpProLysSerGlyThrThr 60
Qy 198 TTGCTGAGGAGGTGGTCTACTTGGTGAGCCAGGCGCTGACCCCGATGATCGGCTTG 257
Db 61 TrpMetHisGluLeuLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
Qy 258 ATGAACATCGACGAGCTCCCGGTCTCGAG-----TACCCA-----CAGCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
Qy 303 GGCGTGGACATCATCAAGAACTGACCTCCCGCTCATCAAGAGCCACTGCCCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
Qy 363 CGCTTTCTGCGCTCTGACCTCCACAATGGAGACTCCCAAGGTTCATCTATATGGCTCGCAAC 422
Db 121 HisLeuLeuProProSerIleTrpLysGluAsnCysIleValTrpValAlaArgAsn 140
Qy 423 CCCAAGATCTGGTGGTCTTATTATCAGTTCCACCGCTCTCTCGCGACCATAGGTAC 482
Db 141 ProlysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACCTTCAAGATTTCTCGCGAGGTATTATGATGATAGCTGGGTACGGCTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlyLys 180
Qy 543 TGGTTTGAGCAGTGCAGGAGTCTCGGAGCAGCCGATGGACTCGAAGCTGCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200
Qy 603 AAGTATGAAGACATCATCGGACCTGGTGACGATGGTGAGCAGCTGGCCAGATTCTGT 662
Db 201 PheTyrGluAspIleLysLysAspProlysArgGluIleGluLysIleLeuLysPheLeu 220
Qy 563 GGGGTGCTGTGACAGGCCAGCTCGAAGCCCTGACGAGCACTGCCACCACTGGTG 722
Db 221 GluLysAspIleSerGluGluIleLeuAsnLysIleIleTyrHisThrSerPheAspVal 240
Qy 723 -----GACCAGTGTGCAACGCTGAGGCCCTGCGCGTGGGC----- 758
Db 241 MetLysGlnAsnProMetThrAsnTyrThrThrLeuProThrSerIleMetAspHisSer 260
Qy 759 -----CGGGGAAGAGTGTGGGTGTGGAGGACATCTTCACCGTCTCC 800
Db 261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAla 280
Qy 801 ATGATAGAAAGTTGACTGGTGTATTAACAGAGATGGGAAAGTGTGACCTCAGCTTT 860
Db 281 GlnAsnGluGluPheAspLysAspTyrGlnLysLysMetAlaGlySerThrLeuThrPhe 300
RESULT 62
ID ABB81792
XX ABB81792 standard; protein; 304 AA.
AC ABB81792;
XX
DT 24-SEP-2002 (first entry)
XX
DE Human sulfotransferase #1.
XX
KW Human; sulfotransferase; cytostatic; gene therapy; cancer; COPD;
chronic obstructive pulmonary disease.
XX
OS Homo sapiens.
XX
PN WO200253713-A2.
XX
PD 11-JUL-2002.
XX
PF 21-DEC-2001; 2001WO-EP015233.
XX
PR 02-JAN-2001; 2001US-0258812P.
XX
PA (FARB ) BAYER AG.
XX
PI Koehler RH;
XX
WI; 2002-528859/56.
XX
N-PSDB; ABB81792, ABB81792.
XX
Novel human sulfotransferase polypeptide useful in therapeutic methods
for treating disorders such as cancer or chronic obstructive pulmonary
disease, and for screening for human sulfotransferase modulators.
XX
Claim 25; Fig 2; 103pp; English.
XX
The invention relates to a novel human sulfotransferase polypeptide. The
sulfotransferase has cytostatic activity. The polynucleotide of the
invention may have a use in gene therapy. The polypeptide and
polynucleotide of the invention are useful in preventing, ameliorating,
or correcting diseases including cancer and chronic obstructive pulmonary
disease (COPD). The sequence represents a human sulfotransferase of the
invention.
XX
Sequence 304 AA;
XX
Alignment Scores:
Pred. No.: 1,45e-30 Length: 304
Score: 422.00 Matches: 97
Percent Similarity: 52.0% Conservative: 59
Best Local Similarity: 32.3% Mismatches: 124
Query Match: 9.6% Indels: 20
DB: Gaps: 5
US-10-768-158-1 (1-2419) x ABB81792 (1-304)
Qy 21 ATGCGGAGAGCGGCGGAGACCCAGACCCCGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy 78 TTCAGTTCATGGGTGGCGTGGCGCCCTTCTGCGCGGGAAGATGGAGGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuLeuSerLysGluTrpTrpGluLysValCys 40
Qy 138 AACTTCCCGTGGCGCGAGCGTGGATCGTACCTACCCAGATCGGCGACCGAGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTrpProLysSerGlyThrThr 60
Qy 198 TTGCTGAGGAGGTGGTCTACTTGGTGAGCCAGGCGCTGACCCCGATGATCGGCTTG 257
Db 61 TrpMetHisGluLeuLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
Qy 258 ATGAACATCGACGAGCTCCCGGTCTCGAG-----TACCCA-----CAGCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
Qy 303 GGCGTGGACATCATCAAGAACTGACCTCCCGCTCATCAAGAGCCACTGCCCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
Qy 363 CGCTTTCTGCGCTCTGACCTCCACAATGGAGACTCCCAAGGTTCATCTATATGGCTCGCAAC 422
Db 121 HisLeuLeuProProSerIleTrpLysGluAsnCysIleValTrpValAlaArgAsn 140
Qy 423 CCCAAGATCTGGTGGTCTTATTATCAGTTCCACCGCTCTCTCGCGACCATAGGTAC 482
Db 141 ProlysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACCTTCAAGATTTCTCGCGAGGTATTATGATGATAGCTGGGTACGGCTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlyLys 180
Qy 543 TGGTTTGAGCAGTGCAGGAGTCTCGGAGCAGCCGATGGACTCGAAGCTGCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200
Qy 603 AAGTATGAAGACATCATCGGACCTGGTGACGATGGTGAGCAGCTGGCCAGATTCTGT 662
Db 201 PheTyrGluAspIleLysLysAspProlysArgGluIleGluLysIleLeuLysPheLeu 220
Qy 563 GGGGTGCTGTGACAGGCCAGCTCGAAGCCCTGACGAGCACTGCCACCACTGGTG 722
Db 221 GluLysAspIleSerGluGluIleLeuAsnLysIleIleTyrHisThrSerPheAspVal 240
Qy 723 -----GACCAGTGTGCAACGCTGAGGCCCTGCGCGTGGGC----- 758
Db 241 MetLysGlnAsnProMetThrAsnTyrThrThrLeuProThrSerIleMetAspHisSer 260
Qy 759 -----CGGGGAAGAGTGTGGGTGTGGAGGACATCTTCACCGTCTCC 800
Db 261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAla 280
Qy 801 ATGATAGAAAGTTGACTGGTGTATTAACAGAGATGGGAAAGTGTGACCTCAGCTTT 860
Db 281 GlnAsnGluGluPheAspLysAspTyrGlnLysLysMetAlaGlySerThrLeuThrPhe 300
```

```
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
Qy 423 CCCAAGATCTGGTGGTCTTATATACAGTTCACCGCTCTCGGACCATAGCTAC 482
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 141 ProLysAspCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAsp 160
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Qy 483 CGAGGCACCTTTCAGAAATCTGCGGAGGTTTATGAATGATAAGCTGGCTACGGCTCC 542
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValIleGlyGlySer 180
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Qy 543 TGGTTTGAACGACGTCAGAGTCTTCGGAGGACCGCAGCTCGAAGCTGCTTTTCTC 602
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 181 TrpPheAspHisValLysGlyTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Qy 603 AAGTATGAACATGCTCGGACCTGCTGACGATGGTGGAGCAGCTGGCAGATCTCTG 662
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 201 PheTyrGluAspIleLysLysAspProLysArgGluIleGluLysIleLeuLysPheLeu 220
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Qy 663 GGGGTGCTCTGTGACAGGCCAGCTGGAAGCCCTGACGGAGCACTGCCACGCTGTG 722
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 221 GluLysAspIleSerGluGluIleLeuAsnLysIleIleTyrHisThrSerPheAspVal 240
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Qy 723 -----GACCAGTGTGCAAGCTGAGGCCCTGCGCCCTGGGCG- 758
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 241 MetLysGlnAsnProMetThrAsnTyrThrThrLeuProThrSerIleMetAspHisSer 260
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Qy 759 -----CGGGGAAGATTGGGCTGGTGGAGGACATCTTCACGCTCTCC 800
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAla 280
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Qy 801 ATGAATGAGAAGTTTGACTGCTGTGTATAACAGAAAGATGGGAAAGTGCTCACGTTT 860
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 281 GlnAsnGluGluPheAspLysAspTyrGlnLysLysMetAlaGlySerThrLeuThrPhe 300
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||

RESULT 63
ADD48894
ID ADD48894 standard; protein; 295 AA.
XX AC ADD48894;
XX AC
XX 02-DEC-2004 (revised)
DT 29-JAN-2004 (first entry)
XX DT
XX
DE Rat Protein P52844, SEQ ID NO 14605.
XX DE
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX KW
XX Rattus norvegicus.
OS Unidentified.
XX OS
XX WO2003016475-A2.
XX PN
XX 27-FEB-2003.
XX PD
XX 14-AUG-2002; 2002WO-US025765.
XX PF
XX 14-AUG-2001; 2001US-0312147P.
XX PR
XX 01-NOV-2001; 2001US-0346382P.
XX PR
XX 26-NOV-2001; 2001US-0333347P.
XX XX
XX (GEO ) GEN HOSPITAL CORP.
XX PA
XX (FARB ) BAYER AG.
XX PA
XX Woolf C, D'urso D, Befort K, Costigan M;
XX PI
XX WPI; 2003-268312/26.
XX DR
XX GENBANK; P52844.
XX DR
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX PT
XX Example 1; Page; 1017pp; English.
XX PS
```

```
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (described in Table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX Cfp.wipo.int/pub/published_pct_sequences.
SQ Sequence 295 AA;
```

```
Alignment Scores:
Pred. No.: 1.77e-30 Length: 295
Score: 421.00 Matches: 91
Percent Similarity: 53.0% Conservative: 60
Best Local Similarity: 31.9% Mismatches: 116
Query Match: 9.6% Indels: 18
DB: Gaps: 3
```

US-10-768-158-1 (1-2419) x ADD48894 (1-295)

```
Qy 60 GAGTTCGAGACCAAGTACTTCGAGTTCATGCGTGGCTGCGGCTTCCTCGCGGGG 119
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 7 GluTyrTyrGluValPheGlyAspPheHisGlyPheLeuMetAspLysArgPheThrLys 26
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Qy 120 AAGTGGAGGAGATCGCAACTTCCCGTGGCGGCCACGACGTGTGGATCGTCACTAC 179
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 27 TyrTrpGluAspIleGluThrPheLeuAlaArgProAspAspLeuLeuValThrTyr 46
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Qy 180 CCCAAGTCCGGCACGAGCTTGTCTGAGAGGTGTCTACTTGGTGAGCCGCGGTGAC 239
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 47 ProLysSerGlySerThrTrpIleSerGluIleValAspMetIleTyrLysGluGlyAsp 66
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Qy 240 CCCGATGAGATCGGCTTGTATGAACATCGACGAGCAGCTCCCGGTCTGTGAGATCCACAG 299
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 67 ValGluLysCysLysGluAspAlaLeuPheAsnArgIleProAspLeuGluCysArgAsn 86
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Qy 300 CCG-----GGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAG 347
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleValLys 106
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Qy 348 AGCCACCTGCGCTACCGCTTTCTGCCCTCTCACCTCCACAATGGAGACTCCAAGGTATC 407
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 107 ThrHisLeuProAlaLysLeuLeuProAlaSerPheTrpGluLysAsnCysLysIleIle 126
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Qy 408 TATATGGCTCGCAACCCCAAGGATCTGGTGTGTCTTATTATCATGTTCCACCGCTCTCTG 467
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 127 TyrLeuCysArgAsnAlaLysAspValValSerTyrTyrTyrTyrPhePheLeuIleMet 146
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Qy 468 CGGACCATGACTACCGGAGGCACCTTTCAGAAATCTCCCGGAGGTTTATGAATGATAG 527
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
Db 147 LysSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGluGlyGln 166
    ||||| :|||:||||| :|||:||||| :|||:||||| :|||:|||||
```



QY 528 CTGGCTACGCTCTCTGTTTGGACACGTGCAGAGAGTTCGGAGACCCGCGATCGACTCG 587  
Db ValProTyrGlySerTrpTyrAspHisValLysSerTrpTyrGluLysSerLysAsnSer 186  
QY 588 AACCTGCTTTTCTCAAGTATGACAGATGATCGGACCTGGTGCAGCATGCTGGAGCAG 647  
Db ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgArgGluValValLys 206  
QY 648 CTGGCCAGATTCCTGGGGGTCTCTGTGACAAAGCCCGACGCTGGAAGCCCTGACGGAGCAC 707  
Db LeuIleGluPheLeuGluArgAspProSerAlaGluLeuValAspArgIleIleGlnHis 226  
QY 708 TGC-----CACCAGCTGGTGCAGCAGCTGCGAACGCTGAGGCCCTGCC----- 752  
Db ThrSerPheGlnGluMetLysAsnAsnProCysThrAsnTyrSerMetLeuProGluThr 246  
QY 753 -----GTGGCCGGGGAGAGTGGCTGTGGAAGGAC 785  
Db MetIleAspLeuLysValSerProPheMetArgLysGlyIleValGlyAspTrpArgAsn 266  
QY 786 ATCTTACCGCTCTCCATGAATGAGAAGTTTGACTTGGTGTATATAACAGAGATGGGAAAG 845  
Db HisPheProGluAlaLeuArgGluArgPheGluGluHisTyrGlnArgHisMetLysAsp 286  
QY 846 TGTGACCTCAGCTTT 860  
Db CysProValLysPhe 291  
RESULT 64  
ABBB1793  
ID ABB81793 standard; protein; 304 AA.  
XX AC ABB81793;  
XX DT 24-SEP-2002 (first entry)  
XX DE Human sulfotransferase #3.  
XX KW Human; sulfotransferase; cytostatic; gene therapy; cancer; COPD;  
XX KW chronic obstructive pulmonary disease.  
XX OS Homo sapiens.  
XX PN WO200253713-A2.  
XX PD 11-JUL-2002.  
XX PF 21-DEC-2001; 2001WO-EP015233.  
XX PR 02-JAN-2001; 2001US-0258812P.  
XX PA (FARB ) BAYER AG.  
XX PI Koehler RH;  
XX WPI; 2002-528859/56.  
XX Novel human sulfotransferase polypeptide useful in therapeutic methods  
PT for treating disorders such as cancer or chronic obstructive pulmonary  
PT disease, and for screening for human sulfotransferase modulators.  
XX Disclosure; Fig 3; 103pp; English.  
XX The invention relates to a novel human sulfotransferase polypeptide. The  
CC sulfotransferase has cytostatic activity. The polynucleotide of the  
CC invention may have a use in gene therapy. The polypeptide and  
CC polynucleotide of the invention are useful in preventing, ameliorating,  
CC or correcting diseases including cancer and chronic obstructive pulmonary  
CC disease (COPD). The sequence represents a human sulfotransferase of the  
XX invention  
XX Sequence 304 AA;

Alignment Scores:  
Pred. No.: 2.5e-30 Length: 304  
Score: 419.50 Matches: 91  
Percent Similarity: 53.0% Conservative: 57  
Best Local Similarity: 32.6% Mismatches: 112  
Query Match: 9.5% Indels: 19  
DB: 5 Gaps: 4  
US-10-768-158-1 (1-2419) x ABB81793 (1-304)  
QY 81 GAGTTCATCGCTGCGGCTGCGCCCTTCTGCGCGGGAAGATGGAGGATCGCCCAAC 140  
Db 22 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 41  
QY 141 TTCCCGTGGCGGCCAGCGACGTGTGTGATCGTACCTACCCCAAGTCCCGCACCAGCTTG 200  
Db 42 PheGlnAlaLysProAspAspLeuLeuIleAlaThrTyrAlaLysAlaGlyThrTrp 61  
QY 201 CTGAGGAGTGTCTACTTGTGTGAGCCAGGGCGCTACCCCGATGAGATCGGCTTGATG 260  
Db 62 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 81  
QY 261 AACATCGACGACGAGCTCCCGGCTCTGGAGTAC-----CCACAGCCG-----GGC 305  
Db 82 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 101  
QY 306 CTGGACATCATCAAGAACTGACCTCTCCCGGCTCATCAAGAGCCACCTGCGCTACCGC 365  
Db 102 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 121  
QY 366 TTTCTGCCCTCTGACCTCCACAATGGAGACTCCAAGGTCTATATATGGTTCGCAACCCC 425  
Db 122 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAsnAla 141  
QY 426 AAGGATCTGTGTGCTTATTATTCAGTTCACCGCTCTCTGCGGACCATGAGCTACCGA 485  
Db 142 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 161  
QY 486 GGCACCTTTCAAGATTCTCGCGAGGTTTATGAATGATAAGCTGGGCTACGGCTCTCGG 545  
Db 162 GlyThrLeuGlyGluTyrIleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 181  
QY 546 TTTGAGCAGCTGACGAGGTTCTGGAGACACCGCATCGAGCTCGAACGCTGTTTCTCAAG 605  
Db 182 TyrAspHisValLysGlyTrpTrpAspValLysAspGlnHisArgIleLeuTyrLeuPhe 201  
QY 606 TATGAGACATGTCATCGGACCTGTGACCATGGTGTGAGCAGCTGCGCAGATTCCTCGGG 665  
Db 202 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 221  
QY 666 GTGCTCTGTGACAAAGCCCGACGCTGGAAGCCCTGACGAGACCTGCCACGAGCTGGT- 722  
Db 222 LysAspIleSerGluGluValLeuAsnLysIleIleTyrHisThrSerPheAspValMet 241  
QY 723 -----GACCAGTGTGCAACGCTGAGGCCCTGCC----- 752  
Db 242 LysGluAsnProMetAlaAsnTyrThrLeuProSerSerIleMetAspHisSerIle 261  
QY 753 -----GTGGCGCGGGAAGAGTGGCTCTGGAAGCAGCATCTTCACCGTCTCCATG 803  
Db 262 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 281  
QY 804 AATGAGAAGTTTCTGCTGGTGTATAAACAAGAACGTGGAAGAGTGCACCTCAGCTTT 860  
Db 282 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 300  
RESULT 65  
AD117047  
ID AD117047 standard; protein; 304 AA.  
XX AC AD117047;  
XX DT 15-APR-2004 (first entry)



QY 366 TTTCTGCCCTCTGACCTCCACATCGAGAGTCCCAAGGTCACTATATATGGCTCGCAACCCC 425  
|||  
Db 122 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAla 141  
|||  
QY 426 AAGGATCTGGTGTCTTATATCAGTTCCACCGCTCTCTCGGACCATGAGCTACCGA 485  
|||  
Db 142 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuLeuProAspPro 161  
|||  
QY 486 GGCACCTTTCAAGAAATTCGCGGAGTTTATGATGATAAGCTGGCTCGGCTCCTCGG 545  
|||  
Db 162 GlyThrLeuGlyGluTyrIleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 181  
|||  
QY 546 TTGAGCAGCTGCAGGAGTTCTGGAGCACCAGCATGGAGCTCGAAGTCTTTTCTCAAG 605  
|||  
Db 182 TyrAspHisValLysGlyTrpTrpAspValLysAspGlnHisArgIleLeuTyrLeuPhe 201  
|||  
QY 606 TATGAAGACATGCATCGGACCTGTGTACGATGGTGGAGCAGCTGGCCAGATTCTCGGG 665  
|||  
Db 202 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 221  
|||  
QY 666 GTGCTCTGTGCAAGCCGAGCTGGAAGCCCTGACGAGCACTGCCACCATGCTGTG--- 722  
|||  
Db 222 LysAspIleSerGluGluValLeuAsnLysIleIleTyrHisThrSerPheAspValMet 241  
|||  
QY 723 -----GACCAGTGTGCAAGCCTGAGCCCTGCCC----- 752  
|||  
Db 242 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 261  
|||  
QY 753 -----GTGGCGGGGAAGAGTGTGGCTGTGGAGGACATCTTCACCGTCTCCATG 803  
|||  
Db 262 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 281  
|||  
QY 804 AATGAGAAGTTTGAGCTTGGTGTATAAACAAGAGATGGGAAAGTGTGACCTCACGTTT 860  
|||  
Db 282 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 300  
|||  
RESULT 66  
ABB81794  
ID ABB81794 standard; protein; 283 AA.  
XX  
AC ABB81794;  
XX  
DT 24-SEP-2002 (first entry)  
XX  
DE Human sulfotransferase #2.  
XX  
KW Human; sulfotransferase; cytostatic; gene therapy; cancer; COPD;  
KW chronic obstructive pulmonary disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200253713-A2.  
XX  
PD 11-JUL-2002.  
XX  
PF 21-DEC-2001; 2001WO-EP015233.  
XX  
PR 02-JAN-2001; 2001US-0258812P.  
XX  
PA (FARB ) BAYER AG.  
XX  
PI Koehler RH;  
XX  
DR WPI; 2002-528859/56.  
DR N-PSDB; ABB83171.  
XX  
PT Novel human sulfotransferase polypeptide useful in therapeutic methods  
PT for treating disorders such as cancer or chronic obstructive pulmonary  
PT disease, and for screening for human sulfotransferase modulators.  
XX  
PS Disclosure; Fig 6; 103pp; English.

CC The invention relates to a novel human sulfotransferase polypeptide The  
CC sulfotransferase has cytostatic activity. The polynucleotide of the  
CC invention may have a use in gene therapy. The polypeptide and  
CC polynucleotide of the invention are useful in preventing, ameliorating,  
CC or correcting diseases including cancer and chronic obstructive pulmonary  
CC disease (COPD). The sequence represents a human sulfotransferase of the  
CC invention  
XX

SQ Sequence 283 AA;

Alignment Scores:  
Pred. No.: 3e-30 Length: 283  
Score: 418.50 Matches: 89  
Percent Similarity: 54.5% Conservative: 55  
Best Local Similarity: 33.7% Mismatches: 101  
Query Match: 9.5% Indels: 19  
DB: 5 Gaps: 4

US-10-768-158-1 (1-2419) x ABB81794 (1-283)

QY 126 GAGGAGATCCGCAACTTCCCGGTGCGGCCGAGCGAGCTGTGGATCGTCACCTACCCCAAG 185  
|||  
Db 16 GluLysValCysAsnPheGlnAlaLysProAspAspLeuIleLeuAlaThrTyrProLys 35  
|||  
QY 186 TCCGGCACCAGCTTGTCTGCGAGGAGTGTCTACTTGTGTAGCCAGGGCGGTGACCCCGAT 245  
|||  
Db 36 SerGlyThrTrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGlu 55  
|||  
QY 246 GAGATCGGCTGTGTAACATCGACGAGCAGCTCCGCTCTGGAG-----TACCCA--- 296  
|||  
Db 56 LysCysLysArgAlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHis 75  
|||  
QY 297 -----CAGCGCGGCGCTGGACATCAAGAACTCTCTCCCGCTCATCAAGAGC 350  
|||  
Db 76 LysGluLysProAspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThr 95  
|||  
QY 351 CACTGCGCTACCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 410  
|||  
Db 96 HisLeuProSerHisLeuIleProSerIleTrpLysGluAsnCysLysIleValTyr 115  
|||  
QY 411 ATGGCTCGCAACCCCAAGGATCTGGTGTGTCTTATTATCAGTTCCACCGCTCTCTGGG 470  
|||  
Db 116 ValAlaArgAsnProLysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSer 135  
|||  
QY 471 ACCATGAGTACCGGAGGACCTTTCAAGAATTTCTGCGGAGGTTTATGAATGATAAGCTG 530  
|||  
Db 136 PheMetProAspProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysVal 155  
|||  
QY 531 GGCTACGGCTCTCGTGGTTGAGCAGCTGCGAGGTTCTGGGAGCACCAGCATGGATCGAAC 590  
|||  
Db 156 ValGlyGlySerTrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArg 175  
|||  
QY 591 GTGCTTTTCTCAAGTATGAGACATCATCGGAGCTGTGGTGTGGTGTGGTGTGGTGTGG 650  
|||  
Db 176 IleLeuTyrLeuPheTyrGluAspIleLysLysAspProLysArgGluIleGluLysIle 195  
|||  
QY 651 GCCAGATTCTCGGGGTGTCTCTGTGACAGCCCGCAGCTGGAAAGCCCTGAGGAGCAGCTGC 710  
|||  
Db 196 LeuLysPheLeuGluLysAspIleSerGluGluIleLeuAsnLysIleIleTyrHisThr 215  
|||  
QY 711 CACGAGTGTGTG-----GACCAGTGTGCAACCGCTGAGCCCTGCGCGTGGGCG--- 758  
|||  
Db 216 SerPheAspValMetLysGlnAsnProMetThrAsnTyrThrThrLeuProThrSerIle 235  
|||  
QY 759 -----CGGGGAAGAGTTGGGCTGTGGAGGAGCATC 788  
|||  
Db 236 MetAspHisSerIleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyr 255  
|||  
QY 789 TTCACTGCTCTCCATGATGAGAAGTTTGACTTGGTGTATTAACAGAGAGATGGGAAAGTGT 848  
|||  
Db 256 PheThrValAlaGlnAsnGluGluPheAspLysAspTyrGlnLysLysMetAlaGlySer 275  
|||  
QY 849 GACCTCACGTTT 860



KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;  
KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;  
KW Alzheimer's disease; infect.

XX Mus musculus.

XX WO2000268649-A2.

XX 06-SEP-2002.

XX 31-JAN-2002; 2002WO-US002785.

XX 31-JAN-2001; 2001US-0265339P.

XX 31-JAN-2001; 2001US-0265412P.

XX 31-JAN-2001; 2001US-0265514P.

XX 02-FEB-2001; 2001US-0265517P.

XX 02-FEB-2001; 2001US-0266406P.

XX 05-FEB-2001; 2001US-0266767P.

XX 07-FEB-2001; 2001US-0266975P.

XX 07-FEB-2001; 2001US-0267057P.

XX 08-FEB-2001; 2001US-0267459P.

XX 09-FEB-2001; 2001US-0267823P.

XX 15-FEB-2001; 2001US-0268974P.

XX 26-FEB-2001; 2001US-0271664P.

XX 27-FEB-2001; 2001US-0271833P.

XX 27-FEB-2001; 2001US-0271855P.

XX 02-MAR-2001; 2001US-0272788P.

XX 02-MAR-2001; 2001US-0273046P.

XX 14-MAR-2001; 2001US-0275925P.

XX 14-MAR-2001; 2001US-0275947P.

XX 14-MAR-2001; 2001US-0275950P.

XX 14-MAR-2001; 2001US-0275989P.

XX 15-MAR-2001; 2001US-0276448P.

XX 15-MAR-2001; 2001US-0276450P.

XX 16-MAR-2001; 2001US-0276397P.

XX 16-MAR-2001; 2001US-0276768P.

XX 20-MAR-2001; 2001US-0278652P.

XX 26-MAR-2001; 2001US-0278775P.

XX 26-MAR-2001; 2001US-0278778P.

XX 29-MAR-2001; 2001US-0279882P.

XX 29-MAR-2001; 2001US-0279884P.

XX 30-MAR-2001; 2001US-0280147P.

XX 11-APR-2001; 2001US-0282992P.

XX 11-APR-2001; 2001US-0283083P.

XX 20-APR-2001; 2001US-0285133P.

XX 23-APR-2001; 2001US-0285749P.

XX 03-MAY-2001; 2001US-0288327P.

XX 03-MAY-2001; 2001US-0288504P.

XX 29-MAY-2001; 2001US-0294047P.

XX 30-MAY-2001; 2001US-0294473P.

XX 08-JUN-2001; 2001US-0296964P.

XX 18-JUN-2001; 2001US-0298959P.

XX 19-JUN-2001; 2001US-0299324P.

XX 13-AUG-2001; 2001US-0312020P.

XX 16-AUG-2001; 2001US-0312889P.

XX 16-AUG-2001; 2001US-0312908P.

XX 21-AUG-2001; 2001US-0313390P.

XX 28-AUG-2001; 2001US-0315470P.

XX 31-AUG-2001; 2001US-0316447P.

XX 07-SEP-2001; 2001US-0318115P.

XX 07-SEP-2001; 2001US-0318118P.

XX 12-SEP-2001; 2001US-0318740P.

XX 19-SEP-2001; 2001US-0323379P.

XX 18-OCT-2001; 2001US-0330245P.

XX 18-OCT-2001; 2001US-0330308P.

XX 14-NOV-2001; 2001US-0332701P.

XX (CURA-) CURAGEN CORP.

XX Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;

XX Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;

XX Gerlach VL, Taupier RU, Gusev VT, Colman SD, Wolenc AR, Pena CEA;

XX Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;

XX

DR WPI; 2002-706998/76.

XX

PT New NOVX polypeptides and nucleic acids, useful for preventing or  
PT treating NOVX-associated disorders, e.g., cancer, cardiomyopathy,  
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.

XX

FS Disclosure; SEQ ID NO 584; 1498pp; English.

XX

This invention relates to a novel nucleic acids, and encoded polypeptides  
thereof, which have properties related to the stimulation of biochemical  
or physiological responses in a cell, tissue, organ or organism.  
Specifically, it refers to the use of biologically active fragments for  
diagnostic and prognostic assays and furthermore in the treatment of  
diverse pathological conditions. The present invention describes novel  
human and murine NOVX proteins, as well as methods to modulate their  
expression using antisense oligos, ribozymes and peptide nucleic acids.  
The NOVX polypeptides, polynucleotides and antibodies are useful in  
treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,  
atherosclerosis, cancer and diabetes. Furthermore, they may be used in  
treating or preventing diseases such as inflammation, autoimmune  
disorders, allergies, blood disorders, acquired immunodeficiency syndrome  
(AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,  
arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy  
and epilepsy. Accordingly, these molecules have many activities including  
cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,  
haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,  
antiasthmatic, nephrotropic, antiarthritic, hepatotropic,  
neuroprotective, nootropic, antibacterial, virucide, antiparasitic,  
relaxant and anticonvulsant. In addition, they are useful in screening  
assays to identify small molecules that modulate or inhibit, for example,  
neurogenesis, wound healing and angiogenesis. The nucleic acids are also  
used as in chromosome mapping, tissue typing, preventive medicine and  
pharmacogenomics. This polypeptide is a homologue of a human NOVX protein  
of the invention.

XX SQ Sequence 304 AA;

Alignment Scores:

Pred. No.: 2,8e-29 Length: 304  
Score: 408.50 Matches: 88  
Percent Similarity: 52.0% Conservative: 57  
Best Local Similarity: 31.5% Mismatches: 115  
Query Match: 9.3% Indels: 19  
DB: 5 Gaps: 3

US-10-768-158-1 (1-2419) x ADI17048 (1-304)

Qy 81 GAGTTCATGGCGGTGCGGCTGCCGCCCTTCTGCCGGGAAGATGGAGAGATGCCCAAC 140

Db 22 GluValAsnGlyIleLeuMetSerLysMetSerGluAsnTrpAspLysIleTrpAsn 41

Qy 141 TTCCCGGTGGCGCCAGCGAGGTGTGGATGTCTACCTACCCAGTCCGGACAGCTTG 200

Db 42 PheGlnAlaLysProAspAspLeuLeuIleAlaThrTyrAlaLysAlaGlyThrTrp 61

Qy 201 CTGCAGGAGGTGGTCTACTTGGTGAGCCAGGGCGCTGACCCCGATGATCGGCTTGATG 260

Db 62 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 81

Qy 261 AACATCGACGACGAGCTCCCGGTCTCTGGAGTACCCACAGCCG-----GGC 305

Db 82 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProProLeuAsnSerGly 101

Qy 306 CTGGACATCATCAAGGAACGTACCTCTCCCGGCTCATCAAGACCGACCTGCTACCGC 365

Db 102 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValGln 121

Qy 366 TTTCTGCCCTCTGACCTCCCAATGGAGACTCCCAAGGTCTATATATGGTTCGCAACCCC 425

Db 122 MetLeuProProSerPheTrpLysGluAsnSerGlnIleIleIleIleValAlaArgAsnAla 141

```
Qy 426 AAGGATCTGGTGTCTATTATTCAGTTCCACCGCTCTCTGCGGACCATGAGCTACCGA 485
Dy 142 LysaspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 161
Qy 486 GGCACCTTTCAGAAATCTGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCCTGG 545
Dy 162 GlyThrLeuGlyGluTyrIleGluThrPheLysAlaGlyLysValLeuTyrGlySerTrp 181
Qy 546 TTGAGCAGCTGCAGGAGTTCTGGGAGCACCAGCTGAGCTGCAAGCTGCTTTTCTCAAG 605
Dy 182 TyrAspHisValLysGlyTyrTrpAspValLysAspLysHisArgIleLeuTyrLeuPhe 201
Qy 606 TATGAAGACATGCATCGGAGCTGCTGACGATGGTGGAGCAGCTGGCCAGATTCTCTGGG 665
Dy 202 TyrGluAspMetLysGluAspProLysArgGluLeuLysLysIleValLysPheLeuGlu 221
Qy 666 GTGCTCTGTGCAAGCCAGCTGGAAGCCCTGAGCGGAGCAGCTGCCACCACTGGTG--- 722
Dy 222 LysAspIleSerGluGluValLeuAsnLysIleIleHisHisThrSerPheAspValMet 241
Qy 723 -----GACCAGTGTGCAAGCTGAGGCCCTGGCC----- 752
Dy 242 LysGlnAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 261
Qy 753 -----GTGGCGGGGAAGAGTTGGGCTGTGGAGGACATCTTCACCGTCTCCATG 803
Dy 262 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 281
Qy 804 AATGAGAAGTTTGACTGGTGTATATAACAGAAAGATGGGAAGTGTGACCTCACGTTT 860
Dy 282 SerGluAspPheAspGluAspTyrArgLysLysMetAlaGlySerThrIleThrPhe 300

RESULT 69
AAE12210
ID AAE12210 standard; protein; 304 AA.
AC AAE12210;
XX
XX 18-DEC-2001 (first entry)
DT Human ST drug-metabolising protein 3 encoded by cDNA.
DE Human; sulfotransferase; drug-metabolising enzyme mediated disorder; ST;
KW DME; toxicity study; gene therapy.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Modified-site 173..175
FT /note= "Protein kinase C phosphorylation site"
FT Modified-site 180..183
FT /note= "Casein kinase II phosphorylation site"
FT Modified-site 194..196
FT /note= "Protein kinase C phosphorylation site"
FT Modified-site 236..239
FT /note= "Casein kinase II phosphorylation site"
FT Modified-site 248..251
FT /note= "Asn is N-glycosylated"
FT Modified-site 258..261
FT /note= "Asn is N-glycosylated"
FT Modified-site 295..300
FT /note= "N-myristoylation site"
XX
XX WO200172977-A2.
PN
XX
XX 04-OCT-2001.
XX
XX 27-MAR-2001; 2001WO-US009659.
XX
XX 27-MAR-2000; 2000US-0192408P.
PR 20-JUN-2000; 2000US-0212725P.
PR 03-JUL-2000; 2000US-00609816.
XX
```

```
PA (PEKE ) PE CORP NY.
XX
XX Woodage T, Wei M, Kodira C, Di Francesco V, Beasley EM;
XX WPI; 2001-616484/71.
DR N-PSDB; AAD19778.
XX
XX New polypeptides useful in labeling reagents with tagged sulfur and
PT identifying modulators of the protein comprises proteins related to
PT sulfotransferase drug-metabolizing enzyme subfamily.
XX
XX Claim 1; Fig 2; 83pp; English.
XX
XX The invention relates to sulfotransferase (ST) drug-metabolising enzyme
CC (DME) peptides, proteins and nucleic acid molecules encoding them.
CC Sequences of the invention are useful in labelling reagents with tagged
CC sulfur, for the development of human therapeutics, as query sequences to
CC perform a search against sequence databases, for identifying other family
CC members or related sequences, to raise antibodies or to elicit another
CC immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, to metabolise
CC compounds for toxicity studies, and as markers for tissues in which the
CC corresponding protein is preferentially expressed. Polypeptides of the
CC invention is also useful in diagnosing a disease, predisposition to a
CC disease mediated by the peptide, in pharmacogenomic analysis and treating
CC a disorder characterised by absence of inappropriate or unwanted
CC expression of the protein. Polynucleotides of the invention are useful
CC for constructing recombinant vectors and host cells, expressing antigenic
CC portions of the proteins, in genotyping and mapping experiments, for
CC designing ribozymes, for monitoring the effectiveness of modulating
CC compounds on the expression or activity of the protein, in diagnostic
CC assays, antisense constructs to control protein expression and in gene
CC therapy. The present sequence is ST drug-metabolising protein 3 encoded
CC by cDNA
XX
XX SQ Sequence 304 AA;
XX
XX Alignment Scores:
XX Pred. No.: 3.9e-29 Length: 304
XX Score: 407.00 Matches: 94
XX Percent Similarity: 51.7% Conservative: 61
XX Best Local Similarity: 31.3% Mismatches: 125
XX Query Match: 9.2% Indels: 20
XX DB: 4 Gaps: 5
XX
XX US-10-768-158-1 (1-2419) x AAE12210 (1-304)
Qy 21 ATGCGGAGGAGCGGCGGAGACCCCGAGACCCCGGGGAGTTCGAG---AGCAAGTAC 77
Dy 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy 78 TTGAGTTCCATGCGTGGCGGCTGCCGCCCTTTCGCCGCGGAGATGGAGGAGATGCC 137
Dy 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTyrTrpGluLysValCys 40
Qy 138 AACTTCCCGTGGCGGCGGAGCGAGCTGTGGATGCTACCTACCCCAAGTCCGGCACCAGC 197
Dy 41 AsnPheGlnAlaLysProAspAspLeuIleLeuAlaThrTyrProLysSerGlyThrThr 60
Qy 198 TTGCTGCAGGAGTGTCTTACTTGTGAGCGGCGGCTGACCCCGCTGATGAGATCGGCTT 257
Dy 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
Qy 258 ATGAACATCGAGCAGCTCCCGGCTCCTCGAG-----TACCCA-----CAGCG 302
Dy 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysGluLysPro 100
Qy 303 GGCTGGACATCATCAAGGAAGTACCTCTCCCGCTCATCAAGACCCACCTCCCTAC 362
Dy 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuLysThrHisLeuProSer 120
Qy 363 CGCTTCTGCCCTCTGACCTCCCAATGGAGTCCCAATGAGGTCTATATGCTCGCAAC 422
```

PA

XX

XX

XX

DR

XX

XX

PT

PT

XX

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

XX

SQ

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-768-158-1 (1-2419) x AAE12210 (1-304)

Qy

Dy

Qy

Dy

Qy

Dy

Qy

Dy

Qy

Dy

Qy

Dy

Qy

Dy

Qy

Dy

Qy

Dy

Qy

```
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
Qy 423 CCCAAGATCTGGTGTCTTATTATCAGTTTCACCGCTCTCTCGGACCATGAGTAC 482
Db 141 ProllysAspCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACCTTCAAGAAATCTCGCGGAGGTTTATGAATGATAGCTGGGTACGGTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlySer 180
Qy 543 TGGTTTGAGCAGCTGCAGAGTCTCGGAGCACCGCATGGACTCGAAGCTCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpIleAlaLysAspThrHisArgIleLeuTyrIleu 200
Qy 603 AAGTATGAAGACATCGCATCGGACCTGTGTACGATGTGTGGACAGCTGGCCAGATTCCTG 662
Db 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy 663 GGGGTGTCCTGTGACAGGCCAGCTGGAAGCCCTGACGGAGCACTGCCACCACTGGTG 722
Db 221 GluLysThrLeuSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
Qy 723 -----GACCAGTGTGCAAGCTGAGGCCCTGCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
Qy 753 -----GTGGCGCGGGAAGAGTTGGGCTGTGGAGGAGCATCTTCACCGTCTCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
Qy 801 ATGAATGAGAAGTTGACTTGGTGTATAAACAGAAAGATGGGAAAGTGAGCTCACGTTT 860
Db 281 MetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300

RESULT 70
ID ABB66006 standard; protein; 316 AA.
XX
AC ABB66006;
DT
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 24810.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX N-PSDB; ABL10109.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
XX
XX Disclosure; SEQ ID NO 24810; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
```

```
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins (ABBS7737-
CC ABBS7072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 316 AA;
Alignment Scores:
Pred. No.: 6.17e-29 Length: 316
Score: 405.00 Matches: 94
Percent Similarity: 53.5% Conservative: 52
Best Local Similarity: 34.4% Mismatches: 99
Query Match: 9.2% Indels: 28
DB: 4 Gaps: 8
US-10-768-158-1 (1-2419) x ABB66006 (1-316)
Qy 126 GAGGAGATCGCAACTTCCCGGTGCGGCCCGAGCGAGCTGTGGATCGTCACCTACCCCAAG 185
Db 45 GluSerIleArgSerLeuProValTyrGlnAspAspValIleMetValSerTyrProArg 64
Qy 186 TCCGGCACCACTGTGCTGTCAGGAGGTGCTCTACTTGTGTGAGCCAG----- 230
Db 65 ThrGlySerThrTrpAlaGlnGluMetValTrpLeuLeuGlyHisGlnLeuAspTyrVal 84
Qy 231 GGGCTGACCCCGAT-----GAGATC-----GGCTTGATGAAC 263
Db 85 AlaAlaGlnAspLeuArgLeuArgSerProLeuIleGluLeuSerAlaLeuPheSer 104
Qy 264 ATGACGAGCAGCTCCCGGTCTGGAGTACCCACAGCCGGCGCTGGAGCATCATCAAGAA 323
Db 105 IleAspHisHisGluThrValAlaGlnLysPheGlyAsnThrValAspLeuValArgAsn 124
Qy 324 CTGACCTCTCCCGCTCATCAAGAGCCACCTGCGCTTACCGCTTCTCTGCCCTCTTGACCTC 383
Db 125 LeuProArgProArgPheAlaArgSerHisLeuProTrpProLeuLeuProGluGlnPhe 144
Qy 384 CACATGGAGACTCCAAGTCACTATATGCTGCGCAACCCCAAGGATCTGTGGTGTCT 443
Db 145 GluThrValLysProArgIleValTyrThrAlaArgAsnProLysAspLeuCysValSer 164
Qy 444 TATTATCAGTTCCACCGCTCTCTCGGACCATGAGTACCGAGGCACCTTTCAAGAATTC 503
Db 165 TyrTyrHisTyrPheLysLeuLeuHisGlyMetAsn-----GlyAspPheGluGlnPhe 182
Qy 504 TGCCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCTCTGGTTTGAGCAGCTGAGAG 563
Db 183 ValAspLeuPheLeuGluGlyHisThrProMetGlySerTyrTrpArgHisValLeuPro 202
Qy 564 TTCTGGGAGCACCGCATGGACTGAGCTGCTTTTCTCAAGTATGAGACATGATCGG 623
Db 203 PheTrpLysArgSerGlnAspAspAsnValLeuPheIleLysTyrGluAspMetValLys 222
Qy 624 GACCTGTGACGATGGTGGAGCAGCTGGCCAGATTCCTGGGGGTGTC-----TGTGAC 677
Db 223 AspLeuProSerValValArgArgCysAlaArgPheLeuGlyValGlnSerLeuLeuAsp 242
Qy 678 AAGGCCAGCTGGAAGCCCTGACGAGCACTGC-----CACCAGCTGGTGGACAG 728
Db 243 ValSerThrLeuGlnLysLeuCysAspHisLeuThrPheAspLysMetArgAlaAsnLys 262
Qy 729 TGCTGCACACCTGAG---GCCCTGCC-----GTGGCGCGGGGAGA 767
Db 263 AlaValAsnLeuGluLysLeuLeuProGluSerSerLysPheIleArgAsnGlyLys 282
Qy 768 GTTGGGCTGTGGAGGACATCTTCCCGTCTCCATGAATGAGAGTTTGACTTGGTGTAT 827
Db 283 IleGlyAspTrpArgAsnHisMetGlyAsnGluMetSerGluArgPheAspGluTrpThr 302
```

CC useful in developmental biology and in elucidating cell signalling and  
cell-cell interactions in higher eukaryotes for the development of  
insecticides, therapeutics and pharmaceutical drugs. The invention  
discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
sequences (AB101840-AB116175) and the encoded proteins (ABBS7737-  
ABBS7072). The sequence data for this patent did not form part of the  
printed specification, but was obtained in electronic format directly  
from WIPO at ftp.wipo.int/pub/published\_pct\_sequences









Pred. No.:	2, 03e-28	Length:	305
Score:	399.50	Matches:	93
Percent Similarity:	51.2%	Conservative:	61
Best Local Similarity:	30.9%	Mismatches:	126
Query Match:	9.1%	Indels:	21
DB:	6	Gaps:	5
US-10-768-158-1 (1-2419) x ABU12099 (1-305)			
Qy	21	ATGCGGAGACGAGCGCGAGACCCCGCGGGAGTTCGAG---AGCAAGTAC	77
Db	1	MetAlaIysThrGluLysAsnAlaProThrMetGluLysLysProGluLysPheAsnIle	20
Qy	78	TTCCAGTTCATGGCGTGGCGTCCGCCCTTCTGCGCGGGAAGATCGGATCGCC	137
Db	21	MetGluValAspGlyValProThrLeuIleLeuSerLysGluIuTrpTrpGluLysValCys	40
Qy	138	AACCTTCCCGTGGCGCGCGAGCGATGTGGATCGTCACCTACCCCAAGTCGGGACACG	197
Db	41	AsnPheGlnAlaLysProAspAspLeuIleAlaThrTyrrProLysSerGlyThrThr	60
Qy	198	TTGCTGAGGAGGTGTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTG	257
Db	61	TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg	80
Qy	258	ATGAACATCGAGCAGCAGCTCCGGTCTCGAG-----TACCCA-----CAGCG	302
Db	81	AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro	100
Qy	303	GGCTTGGACATCATCAAGAACTACCTCCCGCCTCATCAAGAGCCACCTGCCCTAC	362
Db	101	AspLeuGluPheValLeuGluMetSerProGlnLeuIleLysThrHisLeuProSer	120
Qy	363	CGCTTTCGCCCTCTGACCTCCACAAATGGAGACTCCCAAGTTCATCTATATGGCTCGCAAC	422
Db	121	HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrrValAlaArgAsn	140
Qy	423	CCCAAGATCTGGTGTCTTATTATCAGTTTCACCGCTCTCTCGGACCATGAGTAC	482
Db	141	ProLysAspCysLeuValSerTyrrHisPheHisArgMetAlaSerPheMetProAsp	160
Qy	483	CGAGGCACCTTTCAAGAACTTCGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCC	542
Db	161	ProGlnAsnLeuGluGluPheTyrrGluLysPheMetSerGlyLysGlyGluPheGlySer	180
Qy	543	TGTTTTCGACGCTGCAGGAGTTCTGGGAGCAGCCGATGAGCTCGAAGTCTCTTTTCTC	602
Db	181	TrpPheAspHisValLysGlyTrpTrpAlaLysAspMetHisArgIleLeuTyrrLeu	200
Qy	603	AAGTATGAACATGATCGGACCTGGTGCAGTGTGGTGGAGCAGCTGGCCAGATTCTCTG	662
Db	201	PheTyrrGluAspIleLysGlnAsnProLysHisGluIleHisLysValLeuGluPheLeu	220
Qy	663	GGGTGTCTGTGACAGGCCACGTCGAAGCCCTGACGGAGCAGCTGCCACGCTGGTG	722
Db	221	GluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal	240
Qy	723	-----GACAGTGTGCAAGCTGAGCGCTCGCCCTGCGCGTG-----	755
Db	241	MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer	260
Qy	756	-----GGCGGGGAGAGTTGGGCTTGGTTCGGAAGGACATCTTCCCGTC	797
Db	261	IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrVal	280
Qy	798	TCCATGATGAGAGTTTACTTGGTGTATAAACAGAGAGTGGGAAAGTGTGACCTCACG	857
Db	281	AlaLeuAsnGluAsnPheAspLysHisTyrrGluLysLysMetAlaGlySerThrLeuAsn	300
Qy	858	TTT 860	
Db	301	Phe 301	

RESULT 74	
AAE05178	
ID	AAE05178 standard; protein; 304 AA.
XX	
AC	AAE05178;
XX	
DT	12-SEP-2001 (first entry)
XX	
DE	Human drug metabolising enzyme (DME-9) protein.
XX	
KW	Human; drug metabolising enzyme; DME-9; immunosuppressive; gene therapy; cytostatic; autoimmune disorder; inflammatory disorder; atherosclerosis; osteoporosis; eye disorder; hepatic tumour; Addison's disease; cretinism; rheumatoid arthritis; acquired immune deficiency syndrome; AIDS; anaemia; developmental disorder; endocrine disorder; iritis; acromegaly; epilepsy; thyrotoxicosis; pancreatic disorder; diabetes mellitus; obesity; adenoma; gastrointestinal disorder; nodular hyperplasia; conjunctivitis; glaucoma; actinic keratosis; metabolic disorder; dysphagia; anorexia; carcinoma; cell proliferative disorder.
XX	
OS	Homo sapiens.
XX	
FK	Key
FT	Binding-site
FT	Location/Qualifiers
FT	20..304
FT	/label= PAPS_cofactor_binding_site.
XX	
XX	W0200151638-A2.
PN	
XX	
PD	19-JUL-2001.
XX	
PF	12-JAN-2001; 2001WO-US001174.
XX	
PR	14-JAN-2000; 2000US-0176139P.
PR	21-JAN-2000; 2000US-0177443P.
PR	28-JAN-2000; 2000US-0178574P.
XX	
PA	(INCY-) INCYTE GENOMICS INC.
XX	
PI	Yang J, Baughn MR, Burford N, Au-Young J, Lu DAM, Reddy R;
PI	Ring HZ, Hillman JL, Yue H, Azimzai Y, Yao MG, Gandhi AR;
PI	Nguyen DB, Tang YT, Lal P, Bandman O;
XX	
DR	WPI; 2001-425874/45.
DR	N-PSDB; AAD09944.
XX	
PT	Drug metabolizing enzymes and encoding polynucleotides, useful for diagnosing, treating and/or preventing autoimmune, inflammatory, cell proliferative, developmental, endocrine, eye, metabolic, and gastrointestinal disorders.
PT	
PT	
PT	
PS	Claim 1; Page 143-144; 133pp; English.
XX	
CC	The present sequence is human drug metabolising enzyme (DME-9) protein. Human DME and its nucleic acid molecule are useful for the diagnosis, treatment and prevention of disorders associated with increased or decreased expression of DME. Examples of such disorders include, autoimmune/inflammatory disorder such as acquired immune deficiency syndrome (AIDS), rheumatoid arthritis, osteoporosis; cell proliferative disorder such as actinic keratosis, atherosclerosis; developmental disorder such as epilepsy, anaemia; endocrine disorder such as acromegaly, cretinism, thyrotoxicosis; pancreatic disorder such as diabetes mellitus; eye disorder such as conjunctivitis, glaucoma, iritis; metabolic disorder such as Addison's disease, obesity; gastrointestinal disorder such as anorexia, dysphagia and hepatic tumours including nodular hyperplasia, adenomas and carcinomas. DME DNA is useful for creating 'knockin' humanised animals (pigs) or transgenic animals (mice or rats) to model human disease. DME DNA is also in useful in gene therapy. DME and its immunogenic fragments are useful for screening libraries of compounds in several drug screening assays
XX	
SQ	Sequence 304 AA;

Alignment Scores:

Pred. No.:	2.26e-28	Length:	304
Score:	399.00	Matches:	93
Percent Similarity:	51.3%	Conservative:	61
Best Local Similarity:	31.0%	Mismatches:	126
Query Match:	9.1%	Indels:	20
DB:	4	Gaps:	5

  

US-10-768-158-1 (1-2419) x AAE05178 (1-304)	
Qy 21 ATGCGGAGAGCGAGCGAGACCCAGACCCCGGGGGAGTTCGAG---AGCAAGTAC 77	
Db 1 MetAlaIleGluLeuAAsnAlaProThrMetGluLysLysProGluLeuPheAAsnIle 20	
Qy 78 TTCGAGTTCATGCGTGGCGCTCGCGCCCTCTCTCGCGGGAAGATGGAGATCGCC 137	
Db 21 MetGluValAAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys 40	
Qy 138 AACTTCCCGTGGCGCCCGAGCGAGTGTGGATCGTCACCTACCCCAAGTCGGGACACG 197	
Db 41 AsnPheGlnAlaLysProAAspLeuIleLeuAlaThrTyProLysSerGlyThrThr 60	
Qy 198 TTGTCGAGAGGTGTCTACTTGTGTGAGCAGGCGCTGACCCCGATGAGATCGGCTTG 257	
Db 61 TrpMetHisGluIleLeuAAspMetIleLeuAAsnAAspGlyAAspValGluLysCysLysArg 80	
Qy 258 ATGACATCGAGCAGCAGCTCCCGTCTCGGAG-----TACCCA-----CAGCG 302	
Db 81 AlaGlnThrLeuAAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100	
Qy 303 GCGCTGGACATCATCAAGAACTGACCTCTCCCGCCTCATCAAGAGCACCTGCGCTTAC 362	
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120	
Qy 363 CGCTTTCCTCTGACCTCAACAATGAGAGTCCCAAGGTCACTATATATGCTGCGCAAC 422	
Db 121 HisLeuIleProProSerIleTrpLysGluAAsnCysLysIleValTyValAlaArgAAsn 140	
Qy 423 CCAAGGATCGTGGTGTCTTATATCATGTTCCACCGCTCTCGCGACCATGAGCTAC 482	
Db 141 ProLysAAspCysLeuValSerTyTrpHisPheHisArgMetAlaSerPheMetProAAsp 160	
Qy 483 CGAGCAGCTTCAAGAACTTCGCGGAGTTTATGAATGATAAGCTGGCTACGCTCC 542	
Db 161 ProGlnAAsnLeuGluGluPheTyTrpGluLysPheMetSerGlyLysValValGlyArgSer 180	
Qy 543 TGGTTGAGCAGTGCAGAGTCTTGGAGCAGCAGCAGTGCAGTGCAGTCTTTTCTC 602	
Db 181 TrpPheAAspHisValLysGlyTrpTrpAlaAlaLysAAspThrHisArgIleLeuTyLeu 200	
Qy 603 AAGTATGAGACATCATCGGACCTGTGTGACGATGGTGGAGCAGCTGGCCAGATTCCTG 662	
Db 201 PheTyTrpGluAAspIleLysLysAAsnProLysHisGluIleHisValLeuGluPheLeu 220	
Qy 663 GGGGTGCTCTGTGACAAGCCAGCTGGAAGCCCTGACGAGCAGTGCACCAAGCTGGTG 722	
Db 221 GluLysThrLeuSerGlyAAspValIleAAsnLysIleValHisHisThrSerPheAAspVal 240	
Qy 723 -----GACAGTGTGCAACGCTGAGGCGCTCGCC----- 752	
Db 241 MetLysAAspAAsnProMetAlaAAsnHisThrAlaValProAlaHisIlePheAAsnHisSer 260	
Qy 753 -----GTGGCGCGGAGAGTGTGGCTGTGGAGGAGCATCTTCACCGTCTCC 800	
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAAspTrpLysAAsnHisPheThrValAla 280	
Qy 801 ATGAATGAGAAGTTTGAATGCTGTATATAACAGAGATGGGAAAGTGTGACCTCAGCTTT 860	
Db 281 MetAAsnGluAAsnPheAAspLysHisTyTrpGluLysLysMetAlaGlySerThrLeuAAsnPhe 300	

RESULT 75  
ADD12583  
ID ADD12583 standard; protein; 261 AA.  
XX

AC ADD12583;	
XX 01-JAN-2004 (first entry)	
XX Human ENZM-43 protein SEQ ID NO:43.	
XX human; enzyme; ENZM; cytostatic; antiarteriosclerotic; anti-HIV;	
KW antiallergic; cerebroprotective; antiparkinsonian; anticonvulsant;	
KW nootropic; neuroprotective; antiinflammatory; ophthalmological;	
KW antithyroid; antiarthritic; antibacterial; virucide; protozoacide;	
KW antiparasitic; fungicide; anorectic; cardiant; hypotensive;	
KW antifertility; hepatotropic; gene therapy; autoimmune disorder;	
KW inflammatory disorder; AIDS; allergy; atopic dermatitis; arthritis;	
KW thyroiditis; infection; metabolic disorder; obesity;	
KW reproductive disorder; infertility; neurological disorder;	
KW Parkinson's disease; Alzheimer's disease; stroke; epilepsy;	
KW cardiovascular disorder; myocardial infarction; hypertension;	
KW eye disorder; cell proliferative disease; cancer; atherosclerosis;	
hepatitis.	
XX Homo sapiens.	
OS WO2003072729-A2.	
XX 04-SEP-2003.	
XX 21-FEB-2003; 2003WO-US005478.	
XX 22-FEB-2002; 2002US-03659513P.	
PR 19-MAR-2002; 2002US-0365795P.	
XX (INCY-) INCYTE GENOMICS INC.	
XX Sprague WW, Kable AE, Lee S, Griffin JA, Becha SD, Lee EA;	
PI Hafalia AJA, Khare R, Emerling BM, Marquis JP, Ramkumar J;	
PI Elliott VS, Richardson TW, Baughn MR, Jin P, Chien D, Hawkins PR;	
PI Willson AD, Chawla NK, Tran UK, Lee SY, Zebartadian Y, Jiang X;	
PI Jackson AA, Bhatia UG, Burrill JD, Blake JJ, Ho A, Zheng W;	
XX WPI; 2003-731608/69.	
DR N-PSDB; ADD12640.	
XX New human enzymes and polynucleotides, useful for diagnosing, preventing	
PT or treating diseases or conditions associated with aberrant enzyme	
PT expression, e.g. cancer, hepatitis, AIDS, atherosclerosis, obesity,	
XX infection or stroke.	
PS Claim 1; SEQ ID NO 43; 426pp; English.	
XX The present sequence represents a human enzyme designated ENZM-43. Human	
CC enzymes of the present invention have cytostatic, antiarteriosclerotic,	
CC anti-HIV, antiallergic, cerebroprotective, antiparkinsonian,	
CC anticonvulsant, nootropic, neuroprotective, antiinflammatory,	
CC ophthalmological, antithyroid, antiarthritic, antibacterial, virucide,	
CC protozoacide, antiparasitic, fungicide, anorectic, cardiant, hypotensive,	
CC antifertility and hepatotropic activities, and can be used in gene	
CC therapy. The human enzymes and polynucleotides encoding them can be used	
CC in diagnosing, preventing or treating diseases or conditions associated	
CC with the decreased expression or overexpression of the enzymes, such as	
CC autoimmune/inflammatory disorders (AIDS), allergies, atopic dermatitis,	
CC arthritis or thyroiditis), infections (e.g. bacterial, viral, parasitic,	
CC protozoal or fungal), metabolic disorders (e.g. obesity), reproductive	
CC disorders (e.g. infertility), neurological disorders (Parkinson's	
CC disease, Alzheimer's disease, stroke, epilepsy), cardiovascular disorders	
CC (e.g. myocardial infarction or hypertension), eye disorders, or cell	
CC proliferative diseases (e.g. cancer, atherosclerosis, hepatitis). They	
CC are also useful in assessing the effects of exogenous compounds on the	
CC expression of nucleic acid and amino acid sequences of the enzymes.	
XX Sequence 261 AA;	

Alignment Scores: 2.89e-28 Length: 261  
Pred. No.: 261

Score:	397.50	Matches:	85
Percent Similarity:	49.3%	Conservative:	53
Best Local Similarity:	30.4%	Mismatches:	91
Query Match:	9.0%	Indels:	51
DB:	7	Gaps:	3
US-10-768-158-1 (1-2419) x ADD12583 (1-261)			
Qy	72	AAGTACTTCGAGTTCCATGTGGCGTGGCGCTCCGCCCTTCTCCGCGGAAGATGAGGAG	131
Db	12	LysLeuLysGluValGluGlyThrLeuLeuGlnProAlaThrValAspAsnTrpSerGln	31
Qy	132	ATCCCAACTCCCGGTGGCGCCAGGACGTGTGATCGTCACCTACCCCAAGTCGGGC	191
Db	32	IleGlnSerPheGluAlaLysProAspAspLeuLeuIleCysThrTyrProLysAlaGly	51
Qy	192	ACCAGCTTGTCAGGAGGTGCTCTACTTGTGAGCCAGGGCGTGACCCCGATGAGATC	251
Db	52	ThrThrTrpIleGlnGluIleVal-----	59
Qy	252	GGCTTGATGAACATCGACAGCAGCTCCCGGTCTGGAGTACCCACAGCCGGCGCTGGAC	311
Db	60	-----Glu	60
Qy	312	ATCATCAGGAACGTGACCTCTCCCGGCTCATCAGAGCCACCTGCGCTTCTGCTTCTG	371
Db	61	LysAlaLysAlaMetProSerProArgIleLeuLysThrHisLeuSerThrGlnLeuLeu	80
Qy	372	CCCTCTGACCTCCACAATGGAGACTCCAAGTCTATCTATGGCTCGCAACCCCAAGAT	431
Db	81	ProProSerPheTrpGluAsnAsnCysLysPheLeuTyrValAlaAaGAsnAlaLysAsp	100
Qy	432	CTGTGTGTGCTTATTATCATGTTCCACCGCTCTCTGCGGACATCATCCGAGGCACC	491
Db	101	CysMetValSerTyrTyrHisPheGlnArgMetAsnHisMetLeuProAspProGlyThr	120
Qy	492	TTTCAAGAAATCTCCGGAGGTTTATGAATGAATAGCTGGGCTACGGCTCTGTTGTTGAG	551
Db	121	TrpGluGluTyrPheGluThrPheIleAsnGlyLysValValTrpGlySerTrpPheAsp	140
Qy	552	CAGTGTGAGGTTCTGGGAGCACCGATGACCTCGAACGTGCTTTTCTCAAGTATGAA	611
Db	141	HisValLysGlyTyrTrpTrpGluMetCysLysAspArgHisGlnIleLeuPheLysPheTyrGlu	160
Qy	612	GACATGCATCGGACCTGGTGACCATGCTGGAGCAGCTGGCCAGATTCCTGGGGGTGTCC	671
Db	161	AspIleLysArgAspProLysHisGluIleArgLysValMetGlnPheMetGlyLysLys	180
Qy	672	TGTGACAAAGCCCACTGGAAGCCCTGACGGAGCAC-----	707
Db	181	ValAspGluThrValLeuAspLysIleValGlnGlnThrSerPheGluLysMetLysGlu	200
Qy	708	-----TGCACACAGCTGGTGACCACTGTCGCAACGCT	740
Db	201	AsnProMetThrAsnArgSerThrValSerLysSerIleLeuAspGlnSerIleSerSer	220
Qy	741	GAGGCCCTGCGCGTGGCGGGGAAGAGTTGGGCTGTGGAGGACATCTTCACCGTGTCC	800
Db	221	Phe-----MetArgLysGlyThrValGlyAspTrpLysAsnHisPheThrValAla	237
Qy	801	ATGAATGAGAAGTTGATGTTGTATTAACAGAAAGATGGAAAGTGTGACCTCACGTTT	860
Db	238	GlnAsnGluArgPheAspGluIleTyrArgArgLysMetGluGlyThrSerIleAsnPhe	257
RESULT 76			
ADI16670			
ID	ADI16670 standard; protein; 295 AA.		
XX	ADI16670;		
XX	ADI16670;		
DT	15-APR-2004 (first entry)		
XX	Human NOVX protein to treat human pathological conditions SeqID206.		

XX	human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
KW	inflammation; autoimmune disorder; allergy; blood disorder;
KW	acquired immunodeficiency syndrome; AIDS; obesity; asthma;
KW	immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
KW	Alzheimer's disease; infection; stroke; muscular dystrophy; epilepsy;
KW	cytostatic; candidant; antiinflammatory; immunosuppressive; antiallergic;
KW	haemostatic; anti-HIV; antidiabetic; antiarteriosclerotic; anorectic;
KW	antiasthmatic; nephrotropic; antiarthritic; hepatotropic;
KW	neuroprotective; nootropic; antibacterial; virucide; antiparasitic;
KW	relaxant; anticonvulsant; neurogenesis; wound healing; angiogenesis;
KW	chromosome mapping; tissue typing; pharmacogenomic.
XX	
OS	Homo sapiens.
XX	
PN	WO200268649-A2.
XX	
PD	06-SEP-2002.
XX	
PF	31-JAN-2002; 2002WO-US002785.
XX	
PR	31-JAN-2001; 2001US-0265395P.
PR	31-JAN-2001; 2001US-0265412P.
PR	31-JAN-2001; 2001US-0265514P.
PR	31-JAN-2001; 2001US-0265517P.
PR	02-FEB-2001; 2001US-0266406P.
PR	05-FEB-2001; 2001US-0266767P.
PR	07-FEB-2001; 2001US-0266975P.
PR	07-FEB-2001; 2001US-0267057P.
PR	08-FEB-2001; 2001US-0267459P.
PR	09-FEB-2001; 2001US-0267823P.
PR	15-FEB-2001; 2001US-0268974P.
PR	26-FEB-2001; 2001US-0271664P.
PR	27-FEB-2001; 2001US-0271839P.
PR	27-FEB-2001; 2001US-0271855P.
PR	02-MAR-2001; 2001US-0272788P.
PR	02-MAR-2001; 2001US-0273046P.
PR	14-MAR-2001; 2001US-0275925P.
PR	14-MAR-2001; 2001US-0275947P.
PR	14-MAR-2001; 2001US-0275950P.
PR	15-MAR-2001; 2001US-0275989P.
PR	15-MAR-2001; 2001US-0276448P.
PR	16-MAR-2001; 2001US-0276450P.
PR	16-MAR-2001; 2001US-0276397P.
PR	16-MAR-2001; 2001US-0276768P.
PR	20-MAR-2001; 2001US-0278652P.
PR	26-MAR-2001; 2001US-0278775P.
PR	26-MAR-2001; 2001US-0278778P.
PR	29-MAR-2001; 2001US-0279882P.
PR	29-MAR-2001; 2001US-0279884P.
PR	30-MAR-2001; 2001US-0280147P.
PR	11-APR-2001; 2001US-0282992P.
PR	11-APR-2001; 2001US-0283083P.
PR	20-APR-2001; 2001US-0285133P.
PR	23-APR-2001; 2001US-0285749P.
PR	03-MAY-2001; 2001US-0288327P.
PR	03-MAY-2001; 2001US-0288504P.
PR	29-MAY-2001; 2001US-0294047P.
PR	30-MAY-2001; 2001US-0294473P.
PR	08-JUN-2001; 2001US-0296964P.
PR	18-JUN-2001; 2001US-0298959P.
PR	19-JUN-2001; 2001US-0299324P.
PR	13-AUG-2001; 2001US-0312020P.
PR	16-AUG-2001; 2001US-0312889P.
PR	16-AUG-2001; 2001US-0312908P.
PR	21-AUG-2001; 2001US-0313390P.
PR	28-AUG-2001; 2001US-0315470P.
PR	31-AUG-2001; 2001US-0316447P.
PR	07-SEP-2001; 2001US-0318115P.
PR	07-SEP-2001; 2001US-0318118P.
PR	12-SEP-2001; 2001US-0318740P.
PR	19-SEP-2001; 2001US-0323375P.
PR	18-OCT-2001; 2001US-0330245P.

```

PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Tchernev VT, Spytke KA, Zerhusen BD, Patturajan M, Shinkets RA;
XX Li L, Gangolli EA, Radigaru M, Anderson DW, Rastelli L, Miller CE;
XX Gerlach V, Taupier RJ, Gusev V, Colman SD, Wolenc AR, Pena CE;
XX Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
XX
XX WPI; 2002-706998/76.
XX N-PSDB; ADI16669.
XX
XX New NOVX polypeptides and nucleic acids, useful for preventing or
XX treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
XX atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
XX pharmacogenomics.
XX
XX Claim 1; SEQ ID NO 206; 1498pp; English.
XX
XX This invention relates to a novel nucleic acids, and encoded polypeptides
XX thereof, which have properties related to the stimulation of biochemical
XX or physiological responses in a cell, tissue, organ or organism.
XX Specifically, it refers to the use of biologically active fragments for
XX diagnostic and prognostic assays and furthermore in the treatment of
XX diverse pathological conditions. The present invention describes novel
XX human and murine NOVX proteins, as well as methods to modulate their
XX expression using antisense oligos, ribozymes and peptide nucleic acids.
XX The NOVX polypeptides, polynucleotides and antibodies are useful in
XX treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
XX atherosclerosis, cancer and diabetes. Furthermore, they may be used in
XX treating or preventing diseases such as inflammation, autoimmune
XX disorders, allergies, blood disorders, acquired immunodeficiency syndrome
XX (AIDS), obesity, asthma, immunoglobulin (IgA) nephropathy, cirrhosis,
XX arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
XX and epilepsy. Accordingly, these molecules have many activities including
XX cytosstatic, antiinflammatory, immunosuppressive, antiallergic,
XX haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
XX antiaesthetic, nephrotropic, antiarthritic, hepatotropic,
XX neuroprotective, nootropic, antibacterial, viricide, antiparasitic,
XX relaxant and anticonvulsant. In addition, they are useful in screening
XX assays to identify small molecules that modulate or inhibit, for example,
XX neurogenesis, wound healing and angiogenesis. The nucleic acids are also
XX used as in chromosome mapping, tissue typing, preventive medicine and
XX pharmacogenomics. This polypeptide is a human NOVX protein of the
XX invention.
XX
XX Sequence 295 AA;
XX
XX Alignment Scores:
XX Pred. No.: 3.09e-28 Length: 295
XX Score: 397.50 Matches: 85
XX Percent Similarity: 54.2% Conservative: 58
XX Best Local Similarity: 32.2% Mismatches: 102
XX Query Match: 9.0% Indels: 19
XX DB: 5 Gaps: 4
XX
XX US-10-768-158-1 (1-2419) x ADI16670 (1-295)
XX
XX QY 126 GAGGAGATCGCAACTTCCCGGTGCGCCGACGACGAGTGTGAGTGCACCTACCCCAAG 185
XX
XX DB 28 GluLysValCysAsnProAlaLysProAspLeuLeuAlaThrTyrProLys 47
XX
XX QY 186 TCCGGCACCACTGTCTCAGAGGTGTCTACTTGTGTGAGCCAGGGCGCTGACCCCAAT 245
XX
XX DB 48 SerGlyThrThrTyrMetHisGluLeuLeuAspMetIleLeuAsnAspGlyValGlu 67
XX
XX QY 246 GAGATCGCTTGATGAACATCGACGAGCTCCCGGTCTCTGGAG-----TACCCA--- 296
XX
XX DB 68 LysCysLysArgAlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHis 87
XX
XX QY 297 -----CAGCGCGGCTCGGACATCATCAAGGAACCTGACCTCTCCCGGCTCATCAAGAGC 350
XX
XX
XX 88 LysGluLysProAspLeuGluPheValLeuGluMetSerSerProGlnLeuLeuLysThr 107
XX
XX QY 351 CACTGCGCCCTACCGCTTTTGCCTCTGACCTCACAATGAGAGACTCCAAGTCACTAT 410
XX
XX PA ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 108 HisLeuProSerHisLeuLeuProProSerIleTyrPlysGluAenCysLysIleValTyr 127
XX
XX QY 411 ATGGCTCGCAACCCCAAGGATCTGGTGTCTCTTATTATCAGTTCCACCGCTCTCTCGG 470
XX
XX DB 128 ValAlaArgAsnProLysAspCysLeuValSerTyrHisPheHisArgMetAlaSer 147
XX
XX QY 471 ACCATGAGTACCGAGGACCTTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTG 530
XX
XX DB 148 PheMetProAspProGlnAsnLeuGluPheTyrGluLysPheMetSerGlyVal 167
XX
XX QY 531 GGCTACGGCTCTGGTTTGGACGCTGACAGAGTTCTGGAGACCCCATGGAGCTCGAAC 590
XX
XX DB 168 ValGlyGlySerTyrPheAspHisMetLysGlyTyrTrpAlaAlaLysAspMetHisArg 187
XX
XX QY 591 GTGCTTTTCTCAAGTATGAAGACATCATCGGACCTGCTGACGATGGTGGAGCAGCTG 650
XX
XX DB 188 IleLeuTyrLeuPheTyrGluAspIleLysLysAsnProLysHisGluLeuHisVal 207
XX
XX QY 651 GCCAGATTCCTGGGGGTGTCTCTGTGACAAGGCCAGCTGGAAGCCCTGACGAGCAGCTGC 710
XX
XX DB 208 LeuGluPheLeuGluLysThrTyrSerGlyAspValIleAsnLysIleValHisThr 227
XX
XX QY 711 CACGAGTGTGTG-----GACCAGTGTGCACCTGAGGCCCTGCC----- 752
XX
XX DB 228 SerPheAspValMetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIle 247
XX
XX QY 753 -----GTGGCGCGGGAAGAGTTGGGCTGTGGAAGGACATC 788
XX
XX DB 248 PheAsnHisSerIleSerLysPheMetArgLysGlyMetProGlyAspTyrLysAsnHis 267
XX
XX QY 789 TTCACCTCTCCATGAATGAGAGTTTGACTTGTGTATATAACAGAAGATGGGAAAGTGT 848
XX
XX DB 268 PheThrValAlaMetAsnGluAenPheAspLysHisTyrGluLysLysMetAlaGlySer 287
XX
XX QY 849 GACCTCAGCTTT 860
XX
XX DB 288 ThrLeuAsnPhe 291
XX
XX RESULT 77
XX ADN42324
XX ID ADN42324 standard; protein; 295 AA.
XX AC ADN42324;
XX XX
XX DT 17-JUN-2004 (first entry)
XX
XX DE Human novel proteinNOV 56b.
XX
XX KW Human; NOVX; cancer; diabetes; cardiomyopathy; atherosclerosis.
XX
XX OS Homo sapiens.
XX
XX PN US2004033493-A1.
XX
XX PD 19-FEB-2004.
XX
XX PF 31-JAN-2002; 2002US-00072012.
XX
XX PR 31-JAN-2001; 2001US-0265395P.
XX
XX PR 31-JAN-2001; 2001US-0265412P.
XX
XX PR 31-JAN-2001; 2001US-0265514P.
XX
XX PR 02-FEB-2001; 2001US-0265517P.
XX
XX PR 05-FEB-2001; 2001US-0266406P.
XX
XX PR 07-FEB-2001; 2001US-0266767P.
XX
XX PR 07-FEB-2001; 2001US-0266975P.
XX
XX PR 08-FEB-2001; 2001US-0267057P.
XX
XX PR 09-FEB-2001; 2001US-0267459P.
XX
XX PR 15-FEB-2001; 2001US-0268974P.

```

PR 26-FEB-2001; 2001US-0271664P.  
PR 27-FEB-2001; 2001US-0271839P.  
PR 27-FEB-2001; 2001US-0271855P.  
PR 02-MAR-2001; 2001US-0272789P.  
PR 02-MAR-2001; 2001US-0273046P.  
PR 14-MAR-2001; 2001US-0275292P.  
PR 14-MAR-2001; 2001US-0275947P.  
PR 14-MAR-2001; 2001US-0275950P.  
PR 14-MAR-2001; 2001US-0275989P.  
PR 15-MAR-2001; 2001US-0276448P.  
PR 15-MAR-2001; 2001US-0276450P.  
PR 16-MAR-2001; 2001US-0276397P.  
PR 16-MAR-2001; 2001US-0276768P.  
PR 20-MAR-2001; 2001US-0278652P.  
PR 26-MAR-2001; 2001US-0278775P.  
PR 26-MAR-2001; 2001US-0278778P.  
PR 29-MAR-2001; 2001US-0279882P.  
PR 29-MAR-2001; 2001US-0279884P.  
PR 30-MAR-2001; 2001US-0280147P.  
PR 11-APR-2001; 2001US-0282992P.  
PR 11-APR-2001; 2001US-0283083P.  
PR 20-APR-2001; 2001US-0285133P.  
PR 23-APR-2001; 2001US-0285749P.  
PR 03-MAY-2001; 2001US-0288327P.  
PR 03-MAY-2001; 2001US-0288504P.  
PR 29-MAY-2001; 2001US-0294047P.  
PR 30-MAY-2001; 2001US-0294473P.  
PR 08-JUN-2001; 2001US-0296964P.  
PR 18-JUN-2001; 2001US-0298959P.  
PR 19-JUN-2001; 2001US-0299324P.  
PR 13-AUG-2001; 2001US-0312020P.  
PR 16-AUG-2001; 2001US-0312889P.  
PR 16-AUG-2001; 2001US-0312908P.  
PR 21-AUG-2001; 2001US-0313930P.  
PR 28-AUG-2001; 2001US-0315470P.  
PR 31-AUG-2001; 2001US-0316447P.  
PR 07-SEP-2001; 2001US-0318119P.  
PR 07-SEP-2001; 2001US-0318118P.  
PR 12-SEP-2001; 2001US-0318740P.  
PR 19-SEP-2001; 2001US-0323379P.  
PR 18-OCT-2001; 2001US-0330245P.  
PR 18-OCT-2001; 2001US-0330308P.  
PR 14-NOV-2001; 2001US-0332701P.  
XX

(TCHE/) TCHERNEV V T.  
PA (SPYT/) SPYTEK K A.  
PA (ZERH/) ZERHUSEN B D.  
PA (PATT/) PATTURAJAN M.  
PA (SHIM/) SHIMKETS R A.  
PA (LILL/) LI L.  
PA (GANG/) GANGOLLI E A.  
PA (PADI/) PADIGARU M.  
PA (ANDE/) ANDERSON D W.  
PA (RAST/) RASTELLI L.  
PA (MILL/) MILLER C E.  
PA (GERL/) GERLACH V.  
PA (TAUP/) TAUPIER R J.  
PA (GUSE/) GUSEV V Y.  
PA (COLM/) COLMAN S D.  
PA (WOLE/) WOLENC A R.  
PA (PENA/) PENNA C E A.  
PA (FURT/) FURTA K.  
PA (GROS/) GROSSE W M.  
PA (ALSO/) ALSOBROOK J P.  
PA (LEPL/) LEFLEY D M.  
PA (RIEG/) RIEGER D K.  
PA (BURG/) BURGESS C E.  
XX

XX Tchernev VT, Spytke KA, Zerhusen BD, Patturajan M, Shimkets RA;  
PI Li L, Gangolli EA, Padigar M, Anderson DW, Rastelli L, Miller CE,;  
PI Gerlach V, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CE,;  
PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;  
XX

DR WPI; 2004-180039/17.  
XX N-PSDB; ADN42323.  
XX Isolated NOVX polypeptides and polynucleotides, useful for preventing  
PT diagnosing and/or treating cancer, diabetes, cardiomyopathy and  
PT atherosclerosis.  
XX Claim 1; SEQ ID NO 206; 1309pp; English.  
XX The invention relates isolated 162 NOVX polypeptides (NOV1-NOV99,  
CC including splice variants) and the nucleic acids (NA) that encode them.  
CC Also included are the mature NOVX proteins (and their encoding the  
CC polynucleotides), a vector comprising NOVX NA, a cell comprising the  
CC vector, an antibody that binds immunospecifically to NOVX, determining  
CC the presence or amount of NOVX in a sample, identifying an agent that binds to NOVX,  
CC amount of NOVX NA in a sample, identifying an agent that binds to NOVX,  
CC modulating the activity of NOVX, treating or preventing a disease  
CC disorder, determining the presence of or predisposition to a disease  
CC associated with altered levels of NOVX and treating a pathological state  
CC in a mammal comprising administering a polypeptide which is at least 95%  
CC identical to NOVX (or fragment). NOVX and NA may be used in the  
CC prevention, treatment and diagnosis of diseases associated with  
CC inappropriate expression and activity of NOVX (e.g. cancer, diabetes,  
CC cardiomyopathy and/or atherosclerosis). The anti-NOVX antibodies and  
CC antagonists may also be used to down regulate expression and activity of  
CC NOVX. The anti-NOVX antibodies may also be used as diagnostic agents for  
CC detecting the presence of NOVX in samples (e.g. by enzyme linked  
CC immunosorbant assay (ELISA). The agents and methods may be used in this  
CC way to prevent, diagnose and treat cancer, diabetes, cardiomyopathy  
CC and/or atherosclerosis. The present sequence represents a NOVX protein.  
XX

SQ Sequence 295 AA;

Alignment Scores:

Pred. No.:	3,09e-28	Length:	295
Score:	397.50	Matches:	85
Percent Similarity:	54.2%	Conservative:	58
Best Local Similarity:	32.2%	Mismatches:	102
Query Match:	9.0%	Indels:	19
DB:	8	Gaps:	4

US-10-768-158-1 (1-2419) x ADN42324 (1-295)

Qy	126	GAGGAGTCGCCAACTTCCCGTCCGCCAGCAGCGTGTGGATCGTCACCTACCCCAAG	185
Db	28	GlulysValCysAsnPheGlnAlaLysProAspAspLeuLeuAlaThrTyProLys	47
Qy	186	TCCGGCACCAGCTGCTGCAGGAGTGGTCTACTTGTGTGACCCAGGGCGCTGACCCCGAT	245
Db	48	SerGlyThrThrTrpMetHisGluLeuAspMetIleLeuAsnAspGlyAspValGlu	67
Qy	246	GAGATCGGCTTGATGAACATCGACGAGCTCCCGTCTCTGGAG-----TACCACA--	296
Db	68	LysCysLysArgAlaGlnThrLeuAspArgHisAlaPheLeuLeuLysPheProHis	87
Qy	297	-----CAGCGCGCCTGGACATCATCAAGAACTGACCTTCCCGCCTCATCAAGAGC	350
Db	88	LysGluLysProAspLeuGluPheValLeuGluMetSerSerProGlnLeuLysThr	107
Qy	351	CACCTGCCCTACCGCTTCTGCTCTGACCTCCACATGAGAGACTCAAGGTCTCTAT	410
Db	108	HisLeuProSerHisLeuIleProSerIleTrpLysGluAsnCysLysIleValTyr	127
Qy	411	ATGGCTGCAACCCCAAGGATCTGGTGTCTTATTATTCAGTTCCACCGCTCTCTGGGG	470
Db	128	ValAlaArgAsnProLysAspCysLeuValSerTyrTrpHisPheHisArgMetAlaser	147
Qy	471	ACCATGAGTACCGAGGACCTTTCAAGAAATTCCTCCGGAGGTTTATGAATGATAAGCTG	530
Db	148	PheMetProAspProGlnAsnLeuGluPheTyrGluLysPheMetSerGlyLysVal	167
Qy	531	GGCTACGGCTCTCGTTGGTGGACAGCTGCGAGAGTTCTGGGAGCACCCGATCGATCGNAC	590





Db 141 ProlysAspCysLeuValSerTyrTrpHisPheHisArgMetAlaSerPheMetProAsp 160  
Qy 483 CGAGGCACCTTCAAGAAATTCGCGGAGTTTATGAATGATAAGCTGGCTACGGCTCC 542  
Db 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysGlyGluPheGlySer 180  
Qy 543 TGGTTTGAGCAGTCGACGAGTTCCTGGGAGCAGCGCATGGACTCGAAGCTGCTTTTCTC 602  
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200  
Qy 603 AAGTATGAAGACATGCATCGGAGCCTGCTGACGATGGTGGAGCAGCTGGCCAGATTCTCG 662  
Db 201 PheTyrGluAspIleLysGlnAsnProLysHisGluIleHisLysValLeuGluPheLeu 220  
Qy 663 GGGGTGCTCTGACAGGCCAGCTGGAAGCCCTGACGGAGCAGCTGCCACGAGCTGTG 722  
Db 221 GluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240  
Qy 723 -----GACCAAGTGTGCAAGCTGAGCGCTGAGCGCTGCGCCGTG----- 755  
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260  
Qy 756 -----GGCCGGGAGAGCTTGGCTGTGGAAGGACATCTTCACCGTC 797  
Db 261 IleSerLysPheMetArgLysGlyGlyMetProGlyAspTrpLysAsnHisPheThrVal 280  
Qy 798 TCCATGAATGAGAAGTTTGACTTGGTGTATAAACAGAGATGGGAAAGTGTGACCTCACG 857  
Db 281 AlaLeuAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsn 300  
Qy 858 TTT 860  
Db 301 Phe 301  
RESULT 79  
AAE38814  
ID AAE38814 standard; protein; 305 AA.  
XX  
AC AAE38814;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
XX Human sulfotransferase-like protein (POLY13).  
DE Human; tumour; inflammatory disorder; vaccine; gene therapy; cytostatic;  
KW sulfotransferase-like protein; POLY13; chromosome 2.  
KW  
XX Homo sapiens.  
XX  
XX US2003050232-A1.  
XX  
XX 13-MAR-2003.  
XX  
XX 19-APR-2001; 2001US-00839446.  
XX  
XX 19-APR-2000; 2000US-0198293P.  
PR 20-APR-2000; 2000US-0198645P.  
PR 25-APR-2000; 2000US-0198476P.  
PR 26-APR-2000; 2000US-0199880P.  
PR 26-APR-2000; 2000US-0200024P.  
PR 09-JUN-2000; 2000US-0210809P.  
PR 17-JUL-2000; 2000US-0218591P.  
PR 11-AUG-2000; 2000US-0224610P.  
PR 27-FEB-2001; 2001US-0271814P.  
XX  
XX (TAUP/) TAUPIER R J.  
PA (PADI/) PADIGARU M.  
PA (SPVT/) SPYTEK K A.  
PA (BURG/) BURGESS C E.  
PA (VERN/) VERNET C A M.  
PA (FERN/) FERNANDES E R.

PA (SHIM/) SHIMKETS R A.  
PA (LIUX/) LIU X.  
PA (MAJU/) MAJUMDER K.  
PA (COLM/) COLMAN S D.  
PA (ZERH/) ZERHUSEN B D.  
XX  
PI Taupier RJ, Padigaru M, Spytek KA, Burgess CE, Vernet CM;  
PI Fernandes ER, Shimkets RA, Liu X, Majumder K, Colman SD;  
PI Zerhusen BD;  
XX  
DR WPI: 2003-605764/57.  
DR N-PSDB: AAD58965.  
XX  
PT New POLYX nucleic acid, useful for preparing a composition for treating  
PT or preventing e.g., tumor or inflammatory disorder.  
XX  
PS Claim 1; Page 35; 75pp; English.  
XX  
CC The invention relates to new POLYX nucleic acid useful for preparing a  
CC composition for treating or preventing tumour or inflammatory disorder.  
CC The invention is useful as vaccine and in gene therapy. The nucleic acid  
CC is useful for preparing a composition for treating or preventing e.g.,  
CC tumour or inflammatory disorder. The present sequence is human  
CC sulfotransferase-like protein (POLY13). POLY13 gene is located on  
CC chromosome 2. Note: The present sequence is encoded by a DNA containing  
CC translational exceptions  
XX  
SQ Sequence 305 AA;  
Alignment Scores:  
Pred. No.: 3.91e-28 Length: 305  
Score: 396.50 Matches: 93  
Percent Similarity: 50.8% Conservative: 60  
Best Local Similarity: 30.9% Mismatches: 127  
Query Match: 9.0% Indels: 21  
DB: 7 Gaps: 5  
US-10-768-158-1 (1-2419) x AAE38814 (1-305)  
Qy 21 ATGGCGAGAGCGAGCGCGGAGACCCCGAGCAGCCCGGGGAGTTCGAG---AGCAAGTAC 77  
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20  
Qy 78 TTCGAGTTCATGCGGTGCGGCTGCCCGCTTTCGCGCGGAGAGATGGAGAGATCGCC 137  
Db 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys 40  
Qy 138 AACTTCCCGTGGCGGCGGAGCGAGCTGTGGATCGTACCTACCCAGTCCGGCACCAGC 197  
Db 41 AsnPheGlnAlaLysProAspAspLeuIleLeuAlaThrTyrProLysSerGlyThrThr 60  
Qy 198 TTGCTGCAGGAGGTGGTCTACTTGGTGAGCGAGCGGCGCTGACCCCGATGAGATCGGCTG 257  
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80  
Qy 258 ATGAACATCGACGAGCAGCTCCCGGTCTCGAG-----TACCCA-----CAGCCG 302  
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysGluLysPro 100  
Qy 303 GGCTTGACATCATCAGGAAGTACCTCTCCCGCTCATCAGAGCCACCTGCCCTAC 362  
Db 101 AspLeuGluPheValLeuGluMetSerProGlnLeuIleLysThrHisLeuProSer 120  
Qy 363 CGCTTCTGCGCTCTGACCTCCCAATGGAGCTCCCAAGGTCTATCTATATGGCTCGCAAC 422  
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTrpValAlaArgAsn 140  
Qy 423 CCCAAGATCTGGTGGTGTCTTATTATTCAGTTCCACCGCTCTCTCGCGACCATGAGTAC 482  
Db 141 ProlysAspCysLeuValSerTyrTrpHisPheHisArgMetAlaSerPheMetProAsp 160  
Qy 483 CGAGGCACCTTCAAGAAATTCGCGGAGGTTTATGAATGATAGTGGCTACGGCTCC 542

Db 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysGlyGluPheGlySer 180  
 Qy 543 TGGTTTACGACGTCGAGAGTTCTGGAGACCGCATGGACTCAAGCTGCTTTTCTC 602  
 Db 181 TrpPheAspHisValLysGlyTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200  
 Qy 603 AAGTATCAAGACATGTCATCGGACCTGCTGACGATGGTGGAGCAGCTGGCCAGATTCTCTG 662  
 Db 201 PheTyrGluAspIleLysGlnAsnProLysHisGluIleHisLysValLeuGluPheLeu 220  
 Qy 663 GGGGTGCTCTGTGACAAAGCCAGCTGGAAACCCCTGACGGAGCACTGCCACCACTGGTG 722  
 Db 221 GluLysTrpSerGlyAspValIleAsnLysIleValHisHisTrpSerPheAspVal 240  
 Qy 723 -----GACCAAGTCTGCAAGCTGAGCGCTCGCCCTGCGCGTG----- 755  
 Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260  
 Qy 756 -----GGCCGGGGGAAGAGTTGGGCTGTGGAGGACATCTTCACCGTC 797  
 Db 261 IleSerLysPheMetArgLysGlyGlyMetProGlyAspTrpLysAsnHisPheThrVal 280  
 Qy 798 TCATGAATGAGAAAGTTTGACTTGGTGTATTAACAGAGATGGGAAAGTGTGACCTCAGC 857  
 Db 281 AlaLeuAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsn 300  
 Qy 858 TTT 860  
 Db 301 Phe 301  
 RESULT 80  
 ADRI9665  
 ID US-10-768-158-1 (1-2419) x ADRI9665 (1-297)  
 AC ADRI9665;  
 XX  
 DT 07-OCT-2004 (first entry)  
 XX  
 DE Human drug metabolising enzyme (DME) -4 protein sequence.  
 DE  
 KW drug metabolising enzyme; DME; cytostatic; immunosuppressive;  
 KW antiinflammatory; endocrine; ophthalmological; gastrointestinal;  
 KW hepatotropic; cancer; cell proliferative disorder; autoimmune disorder;  
 KW inflammatory disorder; endocrine disorder; eye disorder;  
 KW gastrointestinal disorder; liver disorder; metabolic disorder; enzyme;  
 KW human.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200226998-A2.  
 XX  
 PD 04-APR-2002.  
 XX  
 PF 28-SEP-2001; 2001WO-US030662.  
 XX  
 XX 29-SEP-2000; 2000US-0236947P.  
 PR 06-OCT-2000; 2000US-0238864P.  
 PR 20-OCT-2000; 2000US-0242323P.  
 PR 09-NOV-2000; 2000US-0247581P.  
 PR 16-NOV-2000; 2000US-0249519P.  
 PR 22-NOV-2000; 2000US-0252834P.  
 PR 30-NOV-2000; 2000US-0250567P.  
 XX  
 XX (INCY-) INCYTE GENOMICS INC.  
 XX  
 XX Azimzai Y, Baughin MR, Borowsky ML, Ding L, Duggan BW,  
 PI Elliott VS, Gandhi AR, Griffin JA, Hafalia AJA, Ison CH, Khan FA,  
 PI Lal P, Lee EA, Lu DAM, Nguyen DB, Arvizu C, Policky JL, Ramkumar J;  
 PI Ring HZ, Sanjanwala MS, Tang YT, Tribouley CM, Narinder WK;  
 PI Walsh RT, Warren BA, Xu Y, Yang J, Yao MG, Yue H;  
 XX WPI; 2002-362498/39.  
 DR N-PSDB; ADRI9683.  
 DR

XX  
 PT Human drug metabolizing enzymes, useful in the diagnosis and treatment of  
 PT disorders associated with aberrant (DME) activity, e.g., cancer and  
 PT autoimmune disorders.  
 XX  
 PS Claim 1; SEQ ID NO 4; 142pp; English.  
 XX  
 CC This invention relates to novel drug metabolising enzymes (DME) and the  
 CC nucleotide sequences which encode them. The invention may be useful for  
 CC the development of compounds with a cytostatic, immunosuppressive,  
 CC antiinflammatory, endocrine, ophthalmological, gastrointestinal or  
 CC hepatotropic activity acting as an agonist or antagonist of drug  
 CC metabolising enzyme activity. The invention may be used in the diagnosis  
 CC and treatment of disorders associated with decreased or increased  
 CC expression or activity of drug metabolising enzymes. Such disorders  
 CC include cancer, cell proliferative disorders, autoimmune/inflammatory,  
 CC endocrine, eye, gastrointestinal (including liver disorders) and  
 CC metabolic disorders. The present sequence is that of a human drug  
 CC metabolising enzyme (DME) of the invention. Note: This sequence did not  
 CC form part of the printed specification but was obtained in electronic  
 CC format from EPO.  
 XX  
 SQ Sequence 297 AA;  
 Alignment Scores:  
 Pred. No.: 4.32e-27 Length: 297  
 Score: 385.50 Matches: 84  
 Percent Similarity: 53.7% Conservative: 55  
 Best Local Similarity: 32.4% Mismatches: 105  
 Query Match: 8.8% Indels: 15  
 DB: Gaps: 3  
 US-10-768-158-1 (1-2419) x ADRI9665 (1-297)  
 Qy 126 GAGGAGATCCCAACTTCCCGGTGCGCCGACGAGCTGTGGATCGTCACTACCCCAAG 185  
 Db 36 AspLysIleTrpAsnPheGlnAlaLysProAspAspLeuIleSerThrTyrProLys 55  
 Qy 186 TCCGGACCACTGCTGTCGAGGAGGTGCTACTTGTGTGAGCCAGGGCGGTGACCCCGAT 245  
 Db 56 AlaGlyThrTrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGlu 75  
 Qy 246 GAGATCGGCTGTGATGAATCGACGACGAGCTCCCGGTCTGGAGTAGTACCCACACCGGCG 305  
 Db 76 LysCysLysArgAlaGlnThrLeuAspArgHisAlaPheLeuGluLysPhePro--- 94  
 Qy 306 CTGGACATCATCAAGGAACCTGCTCTCCCGGCTCATCAAGAGCCACTGCTCCCTACCGC 365  
 Db 95 LeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSerHis 114  
 Qy 366 TTTCTGCGCTCTGACCTCCACATGGAGACTCCCAAGGTCTCTCTCGGACCATGAGTACCGA 485  
 Db 115 LeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsnPro 134  
 Qy 426 AAGGATCTGTGTGTCTCTATTATTCAGTTCCACCGCTCTCTCGGACCATGAGTACCGA 485  
 Db 135 LysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSerPheMetProAspPro 154  
 Qy 486 GGCACCTTCAAGAAATTCCTCCGGAGGTTTATGAATGATAGCTGGGTACGGCTCTCTGG 545  
 Db 155 GlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlyArgSerTrp 174  
 Qy 546 TTTGAGCAGCTGCGAGGAGTTCTGGAGCACCGCATGGACTCGAAGCTGCTTTTCTCAAG 605  
 Db 175 PheAspHisValLysGlyTrpTrpAlaLysAspThrHisArgIleLeuTyrLeuPhe 194  
 Qy 606 TATGAAGACATGTCATCGGAGCTGCTGGTACGATGTGGAGCTGGCGAGTCTCTCTGGG 665  
 Db 195 TyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeuGlu 214  
 Qy 666 GTGTCCTGTGACAAAGGCCAGCTGGAAGCCCTGACGAGGACACTGCCACCACTGGTG--- 722  
 Db 215 LysThrLeuSerGlyAspValIleAsnLysIleValHisIleSerPheAspValMet 234

Qy 723 -----GACCAAGTCTGCAACGCTGAGCGCCCTGCCC----- 752  
Db 235 LysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSerIle 254  
Qy 753 -----GTGGCGCGGGAAGAGTTGGCTGTGGAGGACATCTTCACCGTCTCCATG 803  
Db 255 SerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAlaMet 274  
Qy 804 AATGAGAAGTTTGACTTGGTGTATATAACAGAAAGATGGAAAGTGTGACCTCACGTTT 860  
Db 275 AsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 293

RESULT 81  
ABM84008  
ID ABM84008 standard; protein; 262 AA.

AC ABM84008;  
XX 18-NOV-2004 (first entry)  
XX Human diagnostic and therapeutic pprotein SEQ ID NO:4257.  
XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.  
XX Homo sapiens.  
XX WO2004023973-A2.  
XX 25-MAR-2004.  
XX 12-SEP-2003; 2003WO-US028227.  
XX 12-SEP-2002; 2002US-0410259P.  
XX 12-SEP-2002; 2002US-0410260P.  
XX (INCY-) INCYTE CORP.  
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;  
PI Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;  
PI Mooney EM, Deleane AM, Panesar IS, Banville SC, Reddy TP;  
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;  
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LW;  
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;  
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;  
PI Patury S, Shi X, Suarez CJ;  
XX WPI; 2004-329368/30.  
DR N-PSDB; ACN42660.  
XX  
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful  
PT in diagnosing a condition, disease or disorder associated with human  
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or  
PT in gene mapping.  
XX  
XX Claim 27; Page; 190pp; English.  
XX  
XX The invention relates to novel diagnostic and therapeutic polynucleotides  
CC selected from one of the 2722 sequences defined in the specification. A  
CC polynucleotide of the invention may have a use in gene therapy. The human  
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be  
CC used to diagnose a particular condition, disease or disorder associated  
CC with human molecules, e.g. cell proliferative disorders,  
CC autoimmune/inflammatory disorder, developmental disorder, endocrine  
CC disorder, neurological disorders, gastrointestinal disorders, or  
CC infections caused by virus, bacteria, fungi or parasite. The dithp  
CC molecules may also be used in genetic mapping, in identifying individuals  
CC from minute biological samples, in detecting single nucleotide  
CC polymorphisms, as molecular weight markers, and for somatic or germline  
CC gene therapy. The present sequence represents a dithp protein of the  
CC invention. Note: The sequence data for this patent is not represented in  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)

XX SQ Sequence 262 AA;  
Alignment Scores:  
Pred. No.: 5,04e-27 Length: 262  
Score: 384,50 Matches: 84  
Percent Similarity: 48.4% Conservatives: 49  
Best Local Similarity: 30.5% Mismatches: 91  
Query Match: 8.7% Indels: 51  
DB: Gaps: 4  
US-10-768-158-1 (1-2419) x ABM84008 (1-262)  
Qy 90 GCGCTGCGGTGCGCGCCCTTCTGCGCGGGAAGATGAGAGATGCGCAACTTCCCGGTG 149  
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36  
Qy 150 CGGCCAGCAGCGTGTGATCGTCACCTACCCCAAGTCCGCGCACCAGCTTCTCGCAGG 209  
Db 37 ArgProAspAspLeuLeuIleAsnThrTyrProLysSerGlyThrTrpValSerGln 56  
Qy 210 GTGCTCTACTTGTGAGCGCGCTGACCCCGATGAGATCGGCTTGTATGAACATCGAC 269  
Db 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGlyLysCysAsnArgAlaProIleTyr 76  
Qy 270 GAGCAGTCCCGTCTCGGAGTACCCACAGCCGCGCTCGACATCATCAAGAACTGACC 329  
Db 77 ValArgValProPheLeuGluValAsnAspPro----- 87  
Qy 330 TCTCCCGCTCATCAAGAGCCACCTGCCCTTCTGCGCCTCTGACCTCCACAT 389  
Db 87 ----- 87  
Qy 390 GGAGACTCCAAAGGTATCTATATGGCTCGCAACCCCAAGGATCTGGTGGTGTCTTATTAT 449  
Db 88 GlyGluProSerGlyValTyrValAlaArgAsnProLysAspValAlaValSerTyr 107  
Qy 450 CAGTTCCACCGCTCTCTGCGGACCATGAGTCTACCGAGGACCTTTTCAAGAAATTTCCGG 509  
Db 108 HisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAspSerPheLeuGlu 127  
Qy 510 AGGTTTATGAATGATAGCTGGGCTACGGTCTCGCTTGTGAGCAGCTGCGAGGCTTCTGG 569  
Db 128 LysPheMetAlaGlyGluValSerTyrGlySerTrpTyrGlnHisValGlnGluTrpTrp 147  
Qy 570 GAGCACCGCATGGACTCGAAGCTCTTTCTCAAGTATGAAGACATGCATCGCGGACCTG 629  
Db 148 GluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMetLysGluAsnPro 167  
Qy 630 GTGACGATGTGGAGCAGCTGGCCAGATTCTCTGGGGGTGCTCTGTGACAAAGCCGAGCTG 689  
Db 168 LysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuProGluGluThrMet 187  
Qy 690 GAAGCCCTGAGGAGCACTGC----- 710  
Db 188 AspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnProMetThrAsnTyr 207  
Qy 711 -----CACACGCTGGTGACACGAGTGTGCAACGCTGAGGCGCTGCC-----GTG 755  
Db 208 ThrThrValProGlnGluLeuMetAspHisSerIleSer-----ProPheMet 223  
Qy 756 GGCCGGGAAGAGTTGGGCTGTGGAAGGACATCTTCAACCGTCTCCATGAATGAGAAGTTT 815  
Db 224 ArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGlnAsnGluArgPhe 243  
Qy 816 GACTTGTGTATAAACAGAGATGGGAAGTGTGACCTCACGTTT 860  
Db 244 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 258

RESULT 82  
ABM82616  
ID ABM82616 standard; protein; 262 AA.  
XX

```

AC ABM82616;
XX DT 18-NOV-2004 (first entry)
XX DE Human diagnostic and therapeutic pprotein SEQ ID NO:2865.
XX KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX OS Homo sapiens.
XX PN WO2004023973-A2.
XX PD 25-MAR-2004.
XX PF 12-SEP-2003; 2003WO-US028227.
XX PR 12-SEP-2002; 2002US-0410259P.
XX PR 12-SEP-2002; 2002US-0410260P.
XX PA (INCY-) INCYTE CORP.
XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen P;
PI Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Deleane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LI;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
DR WPI: 2004-329368/30.
DR N-PSDB; ACN41268.
XX PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX PS Claim 27; Page: 190pp; English.
XX CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 7222 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC polynucleotide and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX SQ Sequence 262 AA;

Alignment Scores:
Pred. No.: 5,04e-27 Length: 262
Score: 384.50 Matches: 84
Percent Similarity: 48.4% Conservative: 49
Best Local Similarity: 30.5% Mismatches: 91
Query Match: 8.7% Indels: 51
DB: 8 Gaps: 4

US-10-768-158-1 (1-2419) x ABM82616 (1-262)
QY 90 GGCGTGGGTCGCGCCCTTCTGCGGGGAGAGATGGAGGATCGCACTTCCCGTG 149
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36

```













PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226661P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227003P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239933P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0244647P.  
PR 08-NOV-2000; 2000US-0244647P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX WPI; 2001-488787/53.  
DR N-PSDB; AAS30231.  
XX  
PT New polynucleotides and polypeptides, useful for diagnosing, treating, preventing or prognosing e.g. kidney, cardiovascular, blood, electrolyte imbalance or neoplastic disorders, autoimmune diseases, cancers.  
PT  
XX  
PS Claim 1; SEQ ID NO 149; 506pp; English.  
XX  
CC The invention relates to novel nucleic acids and polypeptides useful for diagnosing, treating, preventing and/or prognosing disorders related to these polypeptides. The polynucleotides are especially useful in the diagnosis, prognosis, prevention and/or treatment of diseases which include kidney disorders (e.g. renal failure or nephritis), cardiovascular disorders (e.g. hypertension or myocardial infarction), blood disorders (e.g. anaemia or blood coagulation disorders), electrolyte imbalance disorders (e.g. hyponatraemia or hyperkalaemia), neoplastic disorders (e.g. nephroma or renal cell cancer), autoimmune diseases, cancers, inflammatory diseases, reproductive system disorders, endocrine disorders, neural activity and neurological disorders, wound healing and respiratory disorders. AAU18644-AAU18715 represent the novel human renal and cardiovascular-associated amino acid sequences of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at: ftp.wipo.int/pub/published\_pct\_sequences  
XX





PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0251990P.  
PR 05-JAN-2001; 2000US-0254097P.  
PR 17-JAN-2001; 2001US-0259678P.  
PR 17-JAN-2001; 2001US-00764863.

(ROSE/) ROSEN C A.

(RUBE/) RUBEN S M.

(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-416807/39.

N-PSDB; ACD01486.

New nucleic acids and polypeptides, useful for diagnosing, prognosing, preventing and/or treating e.g. kidney, cardiovascular, blood, electrolyte imbalance, neoplastic, respiratory, or neurological diseases or disorders.

Claim 11; Page 304; 363pp; English.

The invention relates to isolated nucleic acids encoding novel polypeptides. The nucleic acids are useful for chromosome mapping, for radiation hybrid mapping, for detection of cancer, in gene therapy, for identifying individuals from minute biological samples, as an alternative to restriction fragment length polymorphism (RFLP) analysis, in forensic biology and as hybridisation probes for differential identification of tissues or cell types present in a biological sample. Compositions comprising the polynucleotides, polypeptides and antibodies specific for













CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be  
CC used to diagnose a particular condition, disease or disorder associated  
CC with human molecules, e.g. cell proliferative disorders,  
CC autoimmune/inflammatory disorder, developmental disorder, endocrine  
CC disorder, neurological disorders, gastrointestinal disorders, or  
CC infections caused by virus, bacteria, fungi or parasite. The dithp  
CC molecules may also be used in genetic mapping, in identifying individuals  
CC from minute biological samples, in detecting single nucleotide  
CC polymorphisms, as molecular weight markers, and for somatic or germline  
CC gene therapy. The present sequence represents a dithp protein of the  
CC invention. Note: The sequence data for this patent is not represented in  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)  
XX  
XX  
SQ Sequence 254 AA;

Alignment Scores:	2.87e-26	Length:	254
Pred. No.:	Score:	Matches:	83
	376.50	Conservative:	48
Percent Similarity:	47.6%	Mismatches:	85
Best Local Similarity:	30.2%	Indels:	59
Query Match:	6.6%	Gaps:	4
DB:	8		

US-10-768-158-1 (1-2419) x ABM82618 (1-254)

	Qy	90	GGCGTGC	GCGGCTTCGTCCGCCGGAGATGGAAGACATCCCAACTTCCCGG	149
				:::	
Db		17	GlyValProLeuIleLysTyrrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla	36	
				:::	
Qy		150	C GGCC CAGCAGTGTTGATCGTCACTACCACCCCAGTCCGACACCATGTTGCTGCAGGAG	209	
				:::	
Db		37	ArgProAspAsePleuLeuIleAasnThrTyrrProLysSerGlyThrTrpValserGln	56	
				:::	
Qy		210	GTGCTACTTGTGTGAGCCAGGGCGCTCACCCCGATGAGATCGGCTTGATGAACAATCGAC	269	
				:::	
Db		57	IleLeuAspMetIleTyrrGlnGlyaspLeuGluLysCysAsnArgAlaProIleTyrr	76	
				:::	
Qy		270	GAGCAGTCCCGTCTCGAGTAGTACCACAGCGGGCGCTGGACATCATCAAGAAGTA	329	
				:::	
Db		77	ValArgValProPheLeuGlu-----	83	
				:::	
Qy		330	TCTCCCCGCTCATCAAGAGCCACTGCGCTTACC	389	
				:::	
Db		83	-----	----	83
				:::	
Qy		390	GGAGACTCAAAGGTCACTATATGGCTCGCAACCCCAAGGATCTGGTGGTCTATTAT	449	
				:::	
Db		84	-----ValValTyrrValalaargenProLysAspValalavalaserTyrrTyrr	99	
				:::	
Qy		450	CAGTTCACCGCTCTCTCGGACCATTGAGTACCGAGGCACCTTTCAAGAATTTCCCGG	509	
				:::	
Db		100	HisphehisargMetGluLyvalahisProGluProGlyThrTrpAspSerPheLeuGlu	119	
				:::	
Qy		510	AGTTTTATGAATAGATGGCTACGGCTCCTGGTTTGACACGTGAGGAGTTTCGG	569	
				:::	
Db		120	LysPheMetAlaGlyGluValserTyrrGlySerTrpTyrrGlnHisValgInglurTrp	139	
				:::	
Qy		570	GAGCACCGCATGGACTCGAAAGCTGCTTTTCTCAAGTATGAAGACATGCATCGGACCTG	629	
				:::	
Db		140	GluLeuSerArgThrHisProValLeuTyrrLeuPheTyrrGluAspMetLysGluAsnPro	159	
				:::	
Qy		630	GTGACGATGTGGAGCAGCTCGGCAGATTCTCTGGGGGTCTCTGTGACAAGSCCAGCTG	689	
				:::	
Db		160	LysArgGluileGlnLysilLeuGluPheValGlyArgSerLeuProGlugluThrMet	179	
				:::	
Qy		690	GAAGCCCTCAGCGAGCACTGC-----	710	
				:::	
Db		180	AspPheMetValGlnHisThrserrPheLysGluMetLysLysAsnPrometThrAsnTyrr	199	
				:::	
Qy		711	-----CACCAGCTGGTGGAACAGCTGTCGACCGCTAGGCGCCCTGCC---	755	
				:::	
Db		200	ThrThrValProGlnGluLeuMesPhissertileSer-----ProPheMet	215	
				:::	

QY	756	GCCCGGGGAAGAGTTGGCTGTGGAAAGACATCTTCACCGTCTCCATGATGAGAAGTTT	815
Dd	216	ArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGlnAsnGluArgPhe	235
QY	816	CAC TTGGTG TATAAACAG AAGATGGG AAA GTGTG ACCTCAC CGTTT	860
Dd	236	AspAlaaspTyraLaGluLysMetAlaGlyCysSerLeuSerPhe	250
 RESULT_96 ABM82620			
ID	ABM82620 standard; protein; 254 AA.		
XX	AC	ABM82620;	
XX	DT	18-NOV-2004 (first entry)	
XX	DE	Human diagnostic and therapeutic pprotein SEQ ID NO:2869.	
XX	KW	gene therapy; human diagnostic and therapeutic polynucleotide; dithp.	
XX	OS	Homo sapiens.	
XX	FN	WO2004023973-A2.	
XX	PD	25-MAR-2004.	
XX	PF	12-SEP-2003; 2003WO-US028227.	
XX	PR	12-SEP-2002; 2002US-0410259P.	
XX	PA	(INCY-) INCYTE CORP.	

**Alignment Scores:**

```
Pred. No.: 2.87e-26 Length: 254
Score: 376.50 Matches: 83
Percent Similarity: 47.6% Conservative: 48
Best Local Similarity: 30.2% Mismatches: 85
Query Match: 8.6% Indels: 59
DB: 8 Gaps: 4

US-10-768-158-1 (1-2419) x ABM82620 (1-254)
QY 90 GCGTGGCGGCTCCCGCCCTTCTGCGCGGGAAGATGGAGGAGATCGCCAACTTCCCGGTG 149
DB 17 GlyValProLeuIleLeuTyPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
QY 150 CGGCCACGCGGTGGATCGTACCTACCCAGTCGGCAGCCAGCTTCTCGCAGGAG 209
DB 37 ArgProAspLeuLeuIleAsnThrTyrProLysSerGlyThrTrpValSerGln 56
QY 210 GTGGTCTACTTGGTGAGCGAGCGCTGCACCCGATGAGATCGGCTTGATGAACATCGAC 269
DB 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysAsnArgAlaProIleTyr 76
QY 270 GAGCAGCTCCCGGTCTCGAGTACCCACAGCGCGGCTGGACATCATCAAGAACTGACC 329
DB 77 ValArgValProPheLeuGlu----- 83
QY 330 TCTCCCCGCTCATCAAGAGCACCTGCCCTTACCGCTTCTGCGCTTCGACCTCCACAT 389
DB 83 ----- 83
QY 390 GGAGACTCAAGGTCTATATGCTCGCAACCCCAAGGATGCTGGTGTCTTATTAT 449
DB 84 -----ValValTyrValAlaArgAsnProLysAspValAlaValSerTyrTyr 99
QY 450 CAGTTCCACCGCTCTCTCGGACCATGAGTACCGAGCCACCTTTCAGAATTTCTGCCGG 509
DB 100 HisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAspSerPheLeuGlu 119
QY 510 AGGTTTATGATGATAGCTGGCTACCGCTCTGGTTTGACGCTGCAGCTGCAGGATCTGG 569
DB 120 LysPheMetAlaGlyGluValSerTyrGlySerTyrTyrGlnHisValGlnGluTrpTrp 139
QY 570 GAGCACCGCATGAGTCAACGCTGTTTTCTCAAGTATGAAGACATGCATCGGGACCTG 629
DB 140 GluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMetLysGluAsnPro 159
QY 630 GTGACGATGGTGGACAGCTGCCAGATTCCTGGGGGTGTCTCTGACAAAGCCGAGCTG 689
DB 160 LysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuProGluGluThrMet 179
QY 690 GAAGCCCTGACGGAGCAGCTGC----- 710
DB 180 AspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnProMetThrAsnTyr 199
QY 711 -----CACCAGCTGTGACGACGAGTGTCAACGCTGAGCGCCCTGCC--GTG 755
DB 200 ThrThrValProGlnGluLeuMetAspHisSerIleSer-----ProPheMet 215
QY 756 GCGCGGGAAGAGTTGGGTGTGGGAAGGACATCTTCCACCGTCTCCATGAATGAGAAGTTT 815
DB 216 ArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPhe 235
QY 816 GACTTGGTGTATAACAGAAAGTGGAAAGTGTGACCTCACGTTT 860
DB 236 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 250

RESULT 97
ADG76311
ID ADG76311 standard; protein; 259 AA.
XX
AC ADG76311;
XX
DT 11-MAR-2004 (first entry)
XX
```

Human Incyte enzymatic protein (SeqID 3).

enzyme; human; biogenesis; biodegradation; autoimmune; inflammatory; infectious; neurological; cardiovascular; metabolic; cell proliferative; cancer; cytostatic; immunomodulator.

Homo sapiens.

WO2003083082-A2.

09-OCT-2003.

26-MAR-2003; 2003WO-US009772.

29-MAR-2002; 2002US-0368721P.

29-MAR-2002; 2002US-0368799P.

(INCY-) INCYTE CORP.

Tran UK, Kable AE, Ramkumar J, Ison CH, Richardson TW, Lee SY, Khare R, Marquis JP, Swarnakar A, Chawla NK, Elliott VS, Emmerling BM, Becha SD, Hafalia AJA, Li JX, Griffin JA, Hawkins PR, Jin P, Chien D, Jiang X, Jackson AA, Mason PM, Bhatia UG, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W, Gao J;

WPI: 2003-804054/75.

N-PSDB: ADG76366.

New human enzymes (ENZM) polypeptide, useful for preparing a composition for treating a disease associated with decreased expression or overexpression of ENZM e.g. cancer.

Claim 58; SEQ ID NO 3; 385pp; English.

This invention relates to novel isolated polynucleotides, and the enzymes encoded thereof, as well as any biologically active or immunogenic fragments. Specifically, it refers to human enzymes involved in the cellular processes of biogenesis and biodegradation, for example oxidoreductases, hydrolases and isomerases. The present invention describes the use of these nucleic acids and proteins (including single nucleotide polymorphisms) in the diagnosis, treatment and prevention of various disorders such as autoimmune/ inflammatory, infectious, neurological, cardiovascular, metabolic and cell proliferative such as cancer. Furthermore, the polypeptide is useful for preparing a composition for treating a disease associated with decreased expression or overexpression of the functional enzyme (ENZM) e.g. immune disorders. Accordingly, the compositions exhibit cytostatic and immunomodulator activities. This polypeptide sequence is a human Incyte protein, an enzyme of the invention.

Sequence 259 AA;

Alignment Scores:

Pred. No.:	5.04e-25	Length:	259
Score:	363.50	Matches:	83
Percent Similarity:	48.6%	Conservative:	53
Best Local Similarity:	29.8%	Mismatches:	91
Query Match:	8.3%	Indels:	53
DB:	7	Gaps:	5

US-10-768-158-1 (1-2419) x ADG76311 (1-259)

QY 72 AAGTACTTCGAGTTCATGTCGGTGGCTCGCGCTTCTCCGCGGAAGATGGAGGAG 131

DB 12 LysLeuLysGluValGluGlyThrLeuLeuGlnProAlaThrValAspAsnTrpSerGln 31

QY 132 ATCGCCAACTTCCCGTGGCGCCAGGACGCTGGATCTCCTACCTACCCCAAGTCCGCG 191

DB 32 IleGlnSerPheGluAlaLysProAspAspLeuLeuLeuLeuLeuLeuLeuLeuLeu 51

QY 192 ACCAGCTTGTGACGAGGTGGTCTACTTGTGTAGCCAGCGCGCTGACCCCGATGAGATC 251

DB 52 ThrThrTrpIleGlnIleValAspMetIleGluGlnAsnGlyAspValGluLysCys 71



Db 216 ProAsnIlePheThrSerTyrGluArgMetIysGlyGlnLeuGlyGlnValIleSer 235  
Qy 645 CAGCTGCCAGATTCTCTGGGGTGTCTCTGTGACAAAGCCAGCTGGAGCCCTGACGGAG 704  
Db 236 GluValAlaGlnPheLeuGluArgSerValSerValSerGlnGlnMetGlnMetGlnArg 255  
Qy 705 GACTGC-----CACCAGCTGGTGACGAG---TGCTGCAAC----- 737  
Db 256 HisLeuSerPheGluSerMetArgAsnProAlaCysAsnHisValIysGluPheGlu 275  
Qy 738 -----GCTGAGGCCCTGCCCGTGGGGCGGGGAAGA 767  
Db 276 SerMetIysAlaAlaGlyArgGluValGluGluPheArgPheValArgGlyVal 295  
Qy 768 GTTGGGTGTGGAGGACATCTTCACCGTCTCCATGAATGAGAAAGTTTGACTTG 821  
Db 296 ValGlySerHisLysAspGluLeuThrAlaAspIleIleArgGluPheAspLeu 313  
RESULT 99  
AB862612  
ID AB862612 standard; protein; 346 AA.  
XX AC AB862612;  
XX DT 26-MAR-2002 (first entry)  
XX XX Drosophila melanogaster polypeptide SEQ ID NO 14628.  
XX DE Drosophila; developmental biology; cell signalling; insecticide;  
XX KW pharmaceutical.  
XX KW Drosophila melanogaster.  
XX OS WO200171042-A2.  
XX PN 27-SEP-2001.  
XX PD  
XX PF 23-MAR-2001; 2001WO-US009231.  
XX PR 23-MAR-2000; 2000US-0191637P.  
XX PR 11-JUL-2000; 2000US-00614150.  
XX XX (PEKE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX DR N-PSDB; ABL06715.  
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.  
XX PS Disclosure; SEQ ID NO 14628; 21pp + Sequence Listing; English.  
XX CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (AB857737-  
CC AB872072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX XX Sequence 346 AA;  
XX SQ  
Alignment Scores:  
Pred. No.: 1.41e-24 Length: 346  
Score: 359.50 Matches: 82  
Percent Similarity: 46.1% Conservative: 35

Best Local Similarity: 32.3% Mismatches: 88  
Query Match: 8.2% Indels: 49  
DB: 4 Gaps: 5  
US-10-768-158-1 (1-2419) x AB862612 (1-346)  
Qy 84 TTCATGGCGCTGCGG-----CTGCCGCC 107  
Db 27 PheHisGlyGluArgThrGlyPheValGlnValGlySerGluGlyTyrPhePheHis 46  
Qy 108 TTCTGCCCGGGAAGATGGAGAGATCGCAACTTCCTCCGGTGGCGCCAGCGCTGTGG 167  
Db 47 LysTyrLysAspGluAlaGluArgTyrTyrAsnPheGluAlaArgProAspAspValTyr 66  
Qy 168 ATCGTCACCTACCCCAAGTCCGCGACGAGCTTGTCTGTCAGGAGGTGTCTACTTGTGTGAGC 227  
Db 67 IleAlaThrValProArgSerGlyThrTyrTrpThrGlnGluLeuIleTyrLeuValAla 86  
Qy 228 CAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCCAGCAGCTCCCGGCTCTG 287  
Db 87 AsnGlyLeuAspPheGluHisAlaGlnGluArgProLeuThrGluArgPheProPhe 106  
Qy 288 GAGTACCCA----- 296  
Db 107 GluPheProLeuPheValHisProLysIleLysGluGluLeuGlnGluAenArgAsp 126  
Qy 297 -----CAGCGGGCGCTGGACATCATCAAG 320  
Db 127 SerAlaGluAlaLeuGluPheIleGluLysIleAlaArgProGlyTyrGluAlaLeuSer 146  
Qy 321 GAATG-----ACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTCTCTGCC 374  
Db 147 GluIleProArgSerGlnArgPheIleLysThrHisPheProPheSerLeuMetPro 166  
Qy 375 TCTGACCTCCACAATGGAGACTCCAAGTCTATATGCTGCGCAACCCCAAGGATCTG 434  
Db 167 ProSerValLeuGluLysLysCysLysValIleTyrValValArgAspProLysAspVal 186  
Qy 435 GTGGTGTCTTATTATCATGTTCCACCGCTCTCTCGGACCATGAGCTACCGAGGACCTTT 494  
Db 187 AlaValSerTyrTyrHisLeuAsnArgLeuPheArgThrGlnGlyTyrValGlyAspPhe 206  
Qy 495 CAAGAATTCTGCCGAGGTTTATGAATGATAAGCTGGCTACGCTCCTCG----- 545  
Db 207 GluArgTyrTrpHisTyrPheGlnAsn-----GlyLeuAsnProTyrLeuProTyr 223  
Qy 546 TTTGAGCAGCTGCGAGGAGTTCTGGGAGCACCGCATGGACTCGAACGTCGTTTTTCTCAAG 605  
Db 224 TyrSerHisValLysGluAlaArgGluHisAlaHisLeuSerAenValLeuPheLeuArg 243  
Qy 606 TATGAAGACATGATCGGACCTGGACCTGTGACGATGTGGAGCAGCTGGCCAGATTCTCTGGG 665  
Db 244 TyrGluAspMetLeuAlaAspLeuProGlyAlaIleAsnSerIleAlaSerPheLeuGlu 263  
Qy 666 GTGTCTCTGCACAGGCCCGAGCTGGAAAGCCCTGACGAGCAC 707  
Db 264 CysProLysProGluAspMetAspArgLeuLeuAspHis 277  
RESULT 100  
AB862695  
ID AB862695 standard; protein; 254 AA.  
XX AC AB862695;  
XX DT 18-NOV-2004 (first entry)  
XX DE Human diagnostic and therapeutic pprotein SEQ ID NO:2944.  
XX KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.  
XX OS Homo sapiens.  
XX PN WO2004023973-A2.

XX PD 25-MAR-2004.

XX PF 12-SEP-2003; 2003WO-US028227.

XX PR 12-SEP-2002; 2002US-0410259P.

XX PR 12-SEP-2002; 2002US-0410260P.

XX PA (INCY-) INCYTE CORP.

XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F, Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV; Money EM, Deleane AM, Panesar IS, Banville SC, Reddy TP; Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH; Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL; Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES; Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D; Patury S, Shi X, Suarez CJ;

XX WPI; 2004-329368/30.

XX N-PSDB; ACN41347.

XX New diagnostic and therapeutic polynucleotides and polypeptides, useful in diagnosing a condition, disease or disorder associated with human molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or in gene mapping.

XX Claim 27; Page; 190pp; English.

XX The invention relates to novel diagnostic and therapeutic polynucleotides selected from one of the 2722 sequences defined in the specification. A polynucleotide of the invention may have a use in gene therapy. The human diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be used to diagnose a particular condition, disease or disorder associated with human molecules, e.g. cell proliferative disorders, autoimmune/inflammatory disorder, developmental disorder, endocrine disorder, neurological disorders, gastrointestinal disorders, or infections caused by virus, bacteria, fungi or parasite. The dithp molecules may also be used in genetic mapping, in identifying individuals from minute biological samples, in detecting single nucleotide polymorphisms, as molecular weight markers, and for somatic or germline gene therapy. The present sequence represents a dithp protein of the invention. Note: The sequence data for this patent is not represented in the printed specification, but was obtained in electronic format directly from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)

XX SQ Sequence 254 AA;

Alignment Scores:

Pred. No.:	1.86e-24	Length:	254
Score:	357.50	Matches:	81
Percent Similarity:	46.9%	Conservative:	46
Best Local Similarity:	29.9%	Mismatches:	93
Query Match:	8.1%	Indels:	51
DB:	8	Gaps:	4

US-10-768-158-1 (1-2419) x ABM82695 (1-254)

Qy 90 GCGTGGCGCTGCGCCCTTCTGCGGGGAGATGAGGAGATCGCAACTCCCGTG 149

Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36

Qy 150 CGGCCACGACGTGTGGATCGTACCTACCCAGTCCCGCAGCCAGCTTCTGCAGGAG 209

Db 37 ArgProAspAspLeuLeuIleSerThrTyrProLysSerGlyThrTyrValSerGln 56

Qy 210 GTGCTCTACTTGGTACGCGGCGCTGACCCCGATGATGATGCGCTTGATGAACATCGAC 269

Db 57 IleLeuAspMetIleTyrGlnGlyAsp----- 66

Qy 270 GAGCAGCTCCCGGCTCTGGATACCCACAGCGGGCGCTGGACATCATCAAGAACTGACC 329

Db 66 ----- 66

Qy 330 TCTCCCGCCTCATCAAGAGCCACCTGCTTCTGCGCTTCTGCGCTTCTGACCTCCACAAT 389

Db 67 -----LeuGluLysCysHisArg-----ValProGlnThrLeuLeuAsp 79

Qy 390 GGAGACTCCAGGTCATCTATATGGCTCGCAACCCCAAGGATCTGGTGGTCTTATATAT 449

Db 80 GlnLysValLysValTyrValAlaArgAsnAlaLysAspValAlaValSerTyrTyr 99

Qy 450 CAGTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGCCACCTTTCAAGAAATTTGCGCG 509

Db 100 HisPheTyrHisMetAlaLysValHisProGluProGlyThrTyrAspSerPheLeuGlu 119

Qy 510 AGTTTATGAATGATAAGCTGGGCTACGGTCTCTGGTTGAGCACGTCGAGAGTTCTGG 569

Db 120 LysPheMetValGlyGluValSerTyrGlySerTyrPyrGlnHisValGlnGluTrpTrp 139

Qy 570 GAGCACCGCATGGACTCGAACGCTGCTTTTCTCAAGTATGAGACATGCATCGGACCTG 629

Db 140 GluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMetLysGluAsnPro 159

Qy 630 GTGACGATGTGGAGCAGCTGGCCAGATTCTGCGGGGTCTCTGTGACAAGGCCACGCTG 689

Db 160 LysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuProGluGluThrVal 179

Qy 690 GAAGCCCTGACGGAGCAGCTGC-----CACCAGCTGGTGGACCAAGTGTGCAACGCT 740

Db 180 AspPheValValGlnHisThrSerPheLysGluMetLysLysAsnProMetThrAsnTyr 199

Qy 741 GAGCCCTGCGCC-----GTGGGCGGGGAAGA 767

Db 200 ThrThrValProGlnGluPheMetAspHisSerIleSerProPheMetArgLysGlyMet 219

Qy 768 GTTGGGCTGTGGAGGACATCTTCCACGCTCCCATGAATGAGAAGTTTGACTTGGTGTAT 827

Db 220 AlaGlyAspTrpLysThrPheThrValAlaGlnAsnGluArgPheAspAlaAspTyr 239

Qy 828 AAACAGAAAGATGGGAAAGTGTGACCTCACGTTT 860

Db 240 AlaGluLysMetAlaGlyCysSerLeuSerPhe 250

Search completed: May 17, 2006, 09:32:02

Job time : 450.5 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame plus n2p model

Run on: May 17, 2006, 09:18:30 ; Search time 141.6 Seconds  
(without alignments) 4740.705 Million cell u

```
Title: US-10-768-158-1
Perfect score: 4401
Sequence: 1 ggcacggcacgqcgccgc.....aaaaaaaaaaaaaaaaaa 2419
```

Scoring table:	BLOSUM62	
	Xgapop	10.0 , Xgapext 0.5
	Ygapop	10.0 , Ygapext 0.5
	Fgapop	6.0 , Fgapext 7.0
	Delop	6.0 , Delext 7.0

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 5699196

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10
                  Listing first 10

```

```

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/abes/ABSSWEB spool/US10768158/runat_16052006 163422 5282/app_query.fasta_1
-Q=/abes/ABSSWEB spool/US10768158/runat_16052006 163422 5282/app_query.fasta_1
-DB=UniProt -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=100
-DLOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=100 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSExt=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abes02p
-USER=US10768158 @CCN 1 1 466 @runat_16052006 163422 5282 -NCPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THRESHADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

```
Database : Uniprot_7.2.*
          1: uniprot_sprot.*
          2: uniprot_trembl.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1530	34.8	284	1	ST4A1	HUMAN	Q95r01 homo sapien
2	1525	34.7	284	2	Q584X3	PONPY	Q574x3 pongo pygma
3	1508	34.3	284	2	Q865L7	RABIT	Q865l7 oryctolagus
4	1506	34.2	284	1	ST4A1	MOUSE	P63046 mus musculus
5	1506	34.2	284	1	ST4A1	RAT	P63047 rattus norv
6	1506	34.2	284	2	Q37RX5	MOUSE	Q37xy5 mus musculus
7	1483	33.7	284	2	Q8BQV8	MOUSE	Q8bqv8 mus musculus
8	1385	31.5	284	2	Q68EV3	XENLA	Q68ev3 xenopus lae
9	1382	31.4	284	2	Q52KW3	XENLA	Q52kw3 xenopus lae
10	1313	29.8	284	2	Q4SRF8	TEUNG	Q4srf8 tetraodon n
11	1129.5	25.7	218	2	Q5RCA8	PONPY	Q5rca8 pongo pygma
12	495	11.2	298	2	Q9R1S5	MOUSE	Q9r1s5 mus musculus
13	489	11.1	291	1	ST1C1	MOUSE	P52840 mus musculus
14	480	10.9	296	1	ST1C1	MOUSE	Q9G939 mus musculus
15	476	10.8	296	2	Q3UNA9	MOUSE	Q3una9 mus musculus
16	472	10.7	295	1	ST1A3	HUMAN	P50224 h monoamine

90 405.5 9.2 309 2 Q4RG19 TETNG  
 91 405 9.2 316 2 Q9VHHO\_DROME  
 92 404.5 9.2 287 2 Q6DJL6\_XENLA  
 93 402 9.1 287 2 Q5M7C2\_XENLA  
 94 401.5 9.1 293 2 Q49IK6\_BRARE  
 95 401 9.1 304 2 Q6IMIS\_HUMAN  
 96 397.5 9.0 301 1 SF3A1\_RABIT  
 97 395.5 9.0 307 2 Q3SSGA\_HUMAN  
 98 394 8.9 304 2 Q6DJ68\_XENTR  
 99 393.5 8.9 304 2 Q7ZYH0\_XENLA  
 100 392.5 8.9 299 2 Q6GL93\_XENTR

## ALIGNMENTS

## RESULT 1

ST4A1 HUMAN STANDARD; PRT; 284 AA.  
 AC Q9BR01; O43728; integrated into UniProtKB/Swiss-Prot.  
 DT 02-AUG-2002, sequence version 2.  
 DT 07-FEB-2006, entry version 33.  
 DE Sulfotransferase 4A1 (EC 2.8.2.-) (Brain sulfotransferase-like  
 DE protein) (hBR-STL) (hBR-STL-1) (Nervous system sulfotransferase)  
 DE (NST).  
 GN Name=ST4A1; Synonyms=SULTX3;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;

[1] NUCLEOTIDE SEQUENCE (ISOFORM 1), AND CHARACTERIZATION.  
 RP TISSUE=Brain;  
 RC MEDLINE=20164074; PubMed=10698717; DOI=10.1042/0264-6021:3460857;

RX Palany C.N., Xie X., Wang J., Ferrer J., Falany J.L.;  
 RA "Molecular cloning and expression of novel sulphotransferase-like  
 RT cDNAs from human and rat brain."  
 RL Biochem. J. 346:857-864(2000).  
 RN [2]

[3] NUCLEOTIDE SEQUENCE (ISOFORM 1), AND CHARACTERIZATION.  
 RP TISSUE=Brain;  
 RC MEDLINE=22035352; PubMed=12039030; DOI=10.1016/S0378-1119(02)00431-6;

RX Sakakibara Y., Suiko M.C., Pai T.G., Nakayama T., Takami Y.,  
 RA Katsufuchi J., Liu M.-C.;  
 RA "Highly conserved mouse and human brain sulfotransferases: molecular  
 RT cloning, expression, and functional characterization."  
 RL Gene 285:35-47(2002).  
 RN [3]

[4] NUCLEOTIDE SEQUENCE (ISOFORM 1).  
 RP TISSUE=Brain;  
 RC MEDLINE=20164074; PubMed=10698717; DOI=10.1042/0264-6021:3460857;

RX Palany C.N., Xie X., Wang J., Ferrer J., Falany J.L.;  
 RA "Molecular cloning and expression of novel sulphotransferase-like  
 RT cDNAs from human and rat brain."  
 RL Biochem. J. 346:857-864(2000).  
 RN [4]

[5] NUCLEOTIDE SEQUENCE (ISOFORM 1).  
 RP TISSUE=Brain;  
 RC MEDLINE=20164074; PubMed=10698717; DOI=10.1042/0264-6021:3460857;

RX Palany C.N., Xie X., Wang J., Ferrer J., Falany J.L.;  
 RA "Molecular cloning and expression of novel sulphotransferase-like  
 RT cDNAs from human and rat brain."  
 RL Biochem. J. 346:857-864(2000).  
 RN [5]

[6] NUCLEOTIDE SEQUENCE (ISOFORM 2).  
 RP TISSUE=Testis;  
 RC MEDLINE=20164074; PubMed=10698717; DOI=10.1042/0264-6021:3460857;

RX Palany C.N., Xie X., Wang J., Ferrer J., Falany J.L.;  
 RA "Molecular cloning and expression of novel sulphotransferase-like  
 RT cDNAs from human and rat brain."  
 RL Biochem. J. 346:857-864(2000).  
 RN [6]

[7] NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).  
 RP PubMed=15461802; DOI=10.1186/gb-2004-5-10-r84;  
 RX Collins J.E., Wright C.L., Edwards C.A., Davis M.P., Grinham J.A.,  
 RA Cole C.G., Goward M.E., Aguado B., Mallya M., Mokrab Y., Huckle E.J.,  
 RA Beare D.M., Dunham I.,

"A genome annotation-driven approach to cloning the human ORFeome.";  
 Genome Biol. 5:RESEARCH84.1-RESEARCH84.11(2004).  
 [7]

[8] NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RP MEDLINE=20057165; PubMed=10591208; DOI=10.1038/990031;

RX Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,  
 RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,  
 RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,  
 RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,  
 RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,  
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,  
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,  
 RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,  
 RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,  
 RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,  
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,  
 RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,  
 RA Laird G.K., Langford C.F., Leverhulme M.A., Lloyd C., Lloyd D.M.,  
 RA Martyn I.D., Maehreghi-Mohammadi M., Matthews L.H., Mccann O.T.,  
 RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,  
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,  
 RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,  
 RA Scott C.E., Sehara H.K., Skuce C.D., Smalley S., Smith M.L.,  
 RA Soderlund C., Spragon L., Steward C.A., Suleston J.E., Swann R.M.,  
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,  
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,  
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,  
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,  
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,  
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,  
 RA Dorman A., Fang F., Fu Y., Hua A., Kenton S., Lai H., Lao H.I.,  
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,  
 RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,  
 RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,  
 RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,  
 RA Furtos R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,  
 RA Cordun R., Johnson D., Goela D., Graves T., Hawkins J.,  
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,  
 RA Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,  
 RA Korf I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,  
 RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,  
 RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,  
 RA Edelman L., Kim U.J., Shizuya H., Simon M.I., Dumaneki J.P.,  
 RA Peyrard M., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,  
 RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,  
 RA Khan A.S., Lane L., Tilahun Y., Wright H.;  
 RT "The DNA sequence of human chromosome 22."  
 RL Nature 402:489-495(1999).  
 RN [8]

[9] NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).  
 RP TISSUE=Fetal brain, Hippocampus, and Hypothalamus;  
 RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC !- FUNCTION: May catalyze the sulfate conjugation of many drugs,

CC xenobiotic compounds, hormones, and neurotransmitters. Displays  
CC activity towards L-tryptophan, tyrosine, estrone, p-  
CC nitrophenol, 2-naphthylamine, and 2-naphthol. May have a role in the  
CC sulfation of drugs and neurotransmitters in the CNS.  
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q9BR01-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q9BR01-2; Sequence=VSP\_006304;  
CC Note=No experimental confirmation available;  
CC -!- TISSUE SPECIFICITY: Highly expressed in the cerebral cortex and  
CC frontal lobe, slightly less in the cerebellum, occipital and  
CC temporal lobes, relatively low in the medulla and putamen, and  
CC lowest in the spinal cord. No expression detected in the pancreas  
CC (Ref.1). Highly expressed in fetal brain and occipital lobe,  
CC slightly less in the whole brain, frontal lobe, hippocampus, and  
CC lung, very low expression in cerebellum, medulla oblongata,  
CC temporal lobe, testis, kidney and appendix (Ref.2).  
CC -!- SIMILARITY: Belongs to the sulfotransferase family.  
CC  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC  
CC EMBL; AF188698; AAF61197.1; -; mRNA.  
DR EMBL; AF115311; AAF21970.1; -; mRNA.  
DR EMBL; AF176342; AAK64595.1; -; mRNA.  
DR EMBL; AF251263; AAF98152.1; -; mRNA.  
DR EMBL; AL590119; CAC34872.1; -; mRNA.  
DR EMBL; AL456588; CAG30474.1; -; mRNA.  
DR EMBL; Z97055; CAB09788.1; -; Genomic DNA.  
DR EMBL; BC022459; AAH22459.1; -; mRNA.  
DR EMBL; BC028171; AAH28171.1; -; mRNA.  
DR EMBL; BC030665; AAH30665.1; -; mRNA.  
DR HSSP; P50224; 1CJM.  
DR Ensembl; ENSG00000130540; Homo sapiens.  
DR HGNC; HGNC:14903; SULT4A1.  
DR MIM; 608359; gene.  
DR GO; GO:0005829; C:cytosol; NAS.  
DR GO; GO:0008146; F:sulfotransferase activity; NAS.  
DR InterPro; IPR00863; Sulfotransferase.  
DR Pfam; PF00685; Sulfotransferase; 1.  
DR ProDom; PD001218; Sulfotransferase; 1.  
KW Alternative splicing; Lipid metabolism; Steroid metabolism;  
KW Transferase.  
FT CHAIN 1 284 Sulfotransferase 4A1.  
FT FTID=PRO\_0000085167.  
FT NP\_BIND 246 254 FAPS (Potential).  
FT VARSPPLIC 248 284 GRVGLWKDIFTVMNEKFDLVYKQMGKCDLTFDPYL ->  
FT AHCVFARKIFLSW (in isoform 2).  
FT FTID=VSP\_006304.  
FT CONFLICT 55 56 KS -> P (in Ref. 8; AAH30665).  
FT CONFLICT 239 239 N -> S (in Ref. 8; AAH22459).  
SQ SEQUENCE 284 AA; 33085 MW; A6EA6844B66C400B CRC64;  
  
Alignment Scores:  
Pred. No.: 1.17e-99 Length: 284  
Score: 1530.00 Matches: 284  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 34.8% Indels: 0  
DB: 1 Gaps: 0  
  
US-10-768-158-1 (1-2419) x ST4A1\_HUMAN (1-284)  
  
Qy 21 ATGGCGGAGAGCGAGCGGAGACCCCGGCGGGAGTTTCGAGAGCAAGTACTTC 80  
Db 1 MetaLaGluSerGluAlaGluThrProSerThrProGluGluSerLysThrPhe 20  
  
Qy 81 GAGTTCATGCGTGGCGTCCCGCTTCGCGCGGAGAGATGCGGCAATCCCAAC 140  
Db 21 GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluGluAlaAsn 40

Qy 141 TTCCCGGTGCGGCGGCGGAGCGAGCGTGTGGATGTCCTACCTACCCAAAGTCGGGACCAAGCTTG 200  
Db 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60  
  
Qy 201 CTGCAGGAGGTGGTCTACTTCTGGTGAGCCGCGGCTGACCCCGATGAGATCGGCTTGATG 260  
Db 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80  
  
Qy 261 AACATCGACGAGCAGCTCCCGGTCTCTGAGGTACCCACAGCCGGGCTCGACATCATCAAG 320  
Db 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100  
  
Qy 321 GAACGTGACCTCTCCCGGCTCATCAGAGCCACCTGCGCCTTCTGCGCTTCTGAC 380  
Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120  
  
Qy 381 CTCACAAATGAGACTCCCAAGGTATCATATGCTGCTGCAACCCCAAGATCTGCTGGTG 440  
Db 121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuVal 140  
  
Qy 441 TCTTATTATCATGTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGGACCTTTCACGAA 500  
Db 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160  
  
Qy 501 TTCTGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCTCTGTTTGGACACCTGCAG 560  
Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180  
  
Qy 561 GAGTCTCGGAGCACCGCATGAGCTCGAACCTGCTTTTCTCAGATATGAGACATGCGAT 620  
Db 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHis 200  
  
Qy 621 CGGACCTGCTGACGATGCTGGTGGAGCAGCTGCGCCAGATTCTCTGGGGGTCTCTGTGACAAG 680  
Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220  
  
Qy 681 GCCAGCTGGAAAGCCCTGACCGAGCATGCGACAGCTGCTGGTGGACCATGCTGCAACGCT 740  
Db 221 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240  
  
Qy 741 GAGCCCTGCGCGGCGGCGGAGAGTGGGCTGTGGAGGACATCTTCCCGCTCTCC 800  
Db 241 GluAlaLeuProValGlyArgGlyValGlyLeuTrpLysAspIlePheThrValSer 260  
  
Qy 801 ATGAATGAGAAAGTTTACTTGGTGTATATAAACAAGAGATGGGAAAGTGTGACCTCACGTTT 860  
Db 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280  
  
Qy 861 GACTTTTATTTA 872  
Db 281 AspPheTyrLeu 284  
  
RESULT 2  
QSR4X3\_PONPY PRELIMINARY; PRT; 284 AA.  
ID QSR4X3\_PONPY AC QSR4X3;  
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.  
DT 21-DEC-2004, sequence version 1.  
DT 07-FEB-2006, entry version 7.  
DE Hypothetical protein DKFZp459L1114.  
GN Name=DKFZp459L1114;  
OS Pongo pygmaeus (Orangutan).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Pongo.  
OX NCBI\_TaxID=9600;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Cortex;  
RG The German cDNA Consortium;  
RA Bloeker H., Boecher M., Brandt P., Mewes H.W., Weil B., Amid C.,  
RA Osanger A., Fobo G., Han M., Wiemann S.;

```

RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CR861116; CAH93193.1; -; mRNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Hypothetical protein.
SQ SEQUENCE 284 AA; 33111 MW; 531F6C10A3294B00 CRC64;

Alignment Scores:
Pred. No.: 2,65e-99 Length: 284
Score: 1525.00 Matches: 283
Percent Similarity: 99.6% Conservativeness: 0
Best Local Similarity: 99.6% Mismatches: 1
Query Match: 34.7% Indels: 0
DB: Gaps: 0

US-10-768-158-1 (1-2419) x Q865L7_PONPY (1-284)
QY 21 ATGGCGGAGAGCGAGCGGAGACCCCGGCGGAGTTCGAGAGCAAGTACTTC 80
DB 1 MetProGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerIleTyrPhe 20
QY 81 GAGTTCATGGCGTGGCGTCCGCCCTTCTGCGCGGGAAGATCGAGAGATCCCAAC 140
DB 21 GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluGluIleAlaAsn 40
QY 141 TTCCCGGTGGCGCCGAGCGAGCTGGATCGTCACCTACCCCAAGTCCGGACAGCTTG 200
DB 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60
QY 201 CTGCGAGAGGTGGTCTACTTGTGTGAGCAGCGGCGTGCACCCGATGAGATCGGCTTGATG 260
DB 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
QY 261 AACATCGAGAGCAGCTCCCGTCTCTGGAGTACCCACAGCGGCGCTGGACATCATCAAG 320
DB 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100
QY 321 GNACTGACCTCTCCCGGCTCATCAAGAGCCACCTACCGCTTCTGCGCTCTGAC 380
DB 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120
QY 381 CTCACAATGGAGACTCCAAGGTCTATATATGCTCGCAACCCCAAGGATCTGGTGGTG 440
DB 121 LeuHisAsnGlyAspSerIleValIleTyrMetAlaArgAsnProLysAspLeuVal 140
QY 441 TCTTATTATCAGTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGGACCTTTCAAGAA 500
DB 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160
QY 501 TTCTCCCGGAGGTTTATGAATCATAGCTGGCTACGGCTCGCTGCTGTTGAGCACCTGCAG 560
DB 161 PheCysArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180
QY 561 GAGTTCCTGGAGCACCGATGACCTCGAACGTCGCTTTTCTCAAGTATGAGACATGAT 620
DB 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHis 200
QY 621 CGGACCTGGTGCATGGTGGAGCAGCTGGCCAGATTCTCGGGGGTGTCTGTGACCAAG 680
DB 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220
QY 681 GCCAGCTGGAGCCCTGACGAGCAGCTGCCACACGAGCTGGTGACAGCTGCTGCAACGCT 740
DB 221 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240
QY 741 GAGGCGCTCTCCCGTGGCGGGAAGAGTTGGGCTGTGGAAGGACATCTTACCGCTCTCC 800
DB 741 GAGGCGCTCTCCCGTGGCGGGAAGAGTTGGGCTGTGGAAGGACATCTTACCGCTCTCC 800

Db 241 GluAlaLeuProValGlyArgGlyValGlyLeuTrpLysAspIlePheThrValSer 260
QY 801 ATGATGAGAGTTTGGCTTGTATATAACAGAGATGGGAAAGTGTGACCTCACGTTT 860
DB 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280
QY 861 GACTTTTATTATTA 872
DB 281 AspPheTyrLeu 284

RESULT 3
Q865L7_RABIT
ID Q865L7_RABIT PRELIMINARY; PRT; 284 AA.
AC Q865L7_
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE Brain sulfotransferase 4A1.
GN Name=Sult4a1;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA He D., Falany C.N.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY196782; AA045181.1; -; mRNA.
DR HSP; P50224; 1CUM.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 284 AA; 33107 MW; 782B689407A090DE CRC64;

Alignment Scores:
Pred. No.: 4,26e-98 Length: 284
Score: 1508.00 Matches: 279
Percent Similarity: 98.9% Conservativeness: 2
Best Local Similarity: 98.2% Mismatches: 3
Query Match: 34.3% Indels: 0
DB: Gaps: 0

US-10-768-158-1 (1-2419) x Q865L7_RABIT (1-284)
QY 21 ATGGCGGAGAGCGAGCGGAGACCCCGGCGGAGTTCGAGAGCAAGTACTTC 80
DB 1 MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerIleTyrPhe 20
QY 81 GAGTTCATGGTGGTGGCTCGGCTCGGCTCCGCCCTTCTGCGCGGGAAGATCGAGAGATCCCAAC 140
DB 21 GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluGluIleAlaAsn 40
QY 141 TTCCCGGTGGCGCCGAGCGAGCTGGATCGTCACCTACCCCAAGTCCGGACAGCTTG 200
DB 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60
QY 201 CTGCGAGAGGTGGTCTACTTGTGTGAGCAGCGGCGTGCACCCGATGAGATCGGCTTGATG 260
DB 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
QY 261 AACATCGAGAGCAGCTCCCGTCTCTGGAGTACCCACAGCGGCGCTGGACATCATCAAG 320
DB 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100
QY 321 GAACTGACCTCTCCCGGCTCATCAAGAGCACCTGCGCTTCTGCGCTCTGAC 380

```



CC hypothalamus. Not detected in pancreas, liver, lung, intestine,  
 CC kidney, uterus, adrenal gland, thymus, spleen, epididymis,  
 CC testicle, and heart.  
 CC -!- SIMILARITY: Belongs to the sulfotransferase family.  
 CC  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC  
 CC EMBL; AF059257; AAC63999.1; -; mRNA.  
 CC EMBL; AK003034; BAB2522.1; -; mRNA.  
 CC EMBL; AK010293; BAB26829.1; -; mRNA.  
 CC EMBL; BC051132; AAH51132.1; -; mRNA.  
 CC EMBL; BC054757; AAH54757.1; -; mRNA.  
 CC HSP; P50224; ICM.  
 CC MGI; MGI:188971; Sult4a1.  
 CC GO; GO:0005737; C:cytoplasm; TAS.  
 CC GO; GO:0008146; F:sulfotransferase activity; IDA.  
 CC GO; GO:0006790; P:sulfur metabolism; IDA.  
 CC InterPro; IP000863; Sulfotransferase.  
 CC Pfam; PF00685; Sulfotransferase\_1; 1.  
 CC Alternative splicing; Lipid metabolism; Steroid metabolism;  
 KW Transferase.  
 FT CHAIN 1 284 Sulfotransferase 4A1.  
 FT NP\_BIND 246 254 PAPS (Potential).  
 FT VARSP\_LIC 248 284 GRVGLMKDIFVSMNEKEDLVYKQWKGCDLTFDFYL ->  
 FT AHCLFTQIALRWRCRGSGSLRHLCLDLVHVTA (in  
 FT isoform 2).  
 FT /FTID=VSP\_006305.  
 FT CONFLICT 6 6 A -> R (in Ref. 2; BAB2522).  
 FT SEQUENCE 284 AA; 33054 MW; FCAE940F7219B6FC CRC64;  
 Alignment Scores:  
 Pred. No.: 5.91e-98 Length: 284  
 Score: 1506.00 Matches: 278  
 Percent Similarity: 99.3% Conservative: 4  
 Best Local Similarity: 97.9% Mismatches: 2  
 Query Match: 34.2% Indels: 0  
 DB: 1 Gaps: 0  
 US-10-768-158-1 (1-2419) x ST4A1\_MOUSE (1-284)  
 QY 21 ATGGCGGAGAGCGAGCGGAGACCCAGACCCCGGGGGAGTTTCGAGAGCAAGTACTTC 80  
 DB 1 MetAlaGluSerGluAlaGluThrProGlyThrProGlyGluPheGluSerTyrPhe 20  
 QY 81 GAGTTCCATGCGCGTCCGCGTCCGCGCTTCTCCCGCGGAGAGTGGAGAGATCGCCCAAC 140  
 DB 21 GluPheHisGlyValArgLeuProProPheCysArgGlyLysMetGluAspIleAlaAsp 40  
 QY 141 TTCGCGGTCCGCGGAGCGAGTGGATCGTACCTACCCCAAGTCCGCGCACCGCTTG 200  
 DB 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60  
 QY 201 CTGACGAGGAGTGGTCTACTTGGTACGACGCGCGTGCACCCCGATGAGATCGGCTTGATG 260  
 DB 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80  
 QY 261 AACATCGACGACGCTCCGCTCTCGAGTACCCACACCGCGGCTCGACATCATCAAG 320  
 DB 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100  
 QY 321 GAATGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTCTGCCCTCTGAC 380  
 DB 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120  
 QY 381 CTCACATGAGACTCCAAGTCTATATGCTCGCAACCCCAAGATCTGGTGGTG 440  
 DB 121 LeuHisAsnGlyAspSerIleValIleTyrMetAlaArgAsnProLysAspLeuValVal 140  
 QY 441 TCTTATTATCATGTTCCACCGCTCTCTCGCGACCATGAGCTACCGAGGACCTTTCAGAA 500  
 DB 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160

QY 501 TTCTGCGGAGGTTTATGAATGATAAGCTGGGTACGGCTCCTGGTTTGACACGTGCAG 560  
 DB 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180  
 QY 561 GAGTTCTGGGAGCACCGCTGACTCGAACTGCTTTTCTCAAGTATGAAGACATGCAT 620  
 DB 181 GluPheTrpGluHisArgMetAspAlaAsnValLeuPheLeuLysTyrGluAspMetHis 200  
 QY 621 CGGACCTGGTGACGATGGTGGACAGCTGGCCAGATTCCTGGGGGTGCTCTGTGACAAG 680  
 DB 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220  
 QY 681 GCCAGCTGGAGCCCTGACGGAGCACTGCCACAGCTGGTGGACCACTGCTGCAACGCT 740  
 DB 221 AlaGlnLeuGluSerLeuIleGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240  
 QY 741 GAGGCCCTGCCCGTGGCGCGGGAAGTAGTGGGCTGTGGAAGGACATCTTCACCGTCTCC 800  
 DB 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 260  
 QY 801 ATGAATGAGAAGTTTGAATCTGCTGTATAAACAGAGATGGGAAAGTGGAAAGTGTGACCTT 860  
 DB 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280  
 QY 861 GACTTTTATTTA 872  
 DB 281 AspPheTyrLeu 284  
 RESULT 5  
 ST4A1\_RAT STANDARD; PRT; 284 AA.  
 ID P63047; O88872; Q91XS5; Q9CW7; Q9DC97;  
 DT 31-AUG-2004, integrated into UniProtKB/Swiss-Prot.  
 DT 31-AUG-2004, sequence version 1.  
 DT 07-FEB-2006, entry version 14.  
 DE Sulfotransferase 4A1 (EC 2.8.2.-) (Brain sulfotransferase-like  
 DE protein) (rBR-STL) (Nervous system sulfotransferase) (NST).  
 GN Name=Sult4a1; Synonyms=Sultx3;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [MRNA], TISSUE SPECIFICITY, DEVELOPMENTAL STAGE,  
 RP AND SUBCELLULAR LOCATION.  
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;  
 RX MEDLINE=20164074; PubMed=10698717; DOI=10.1042/0264-6021.3460857;  
 RA Falany C.N., Xie X., Wang J., Ferrer J., Falany J.L.;  
 RT "Molecular cloning and expression of novel sulphotransferase-like  
 RT cDNAs from human and rat brain."  
 RL Biochem. J. 346:857-864 (2000).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [MRNA] OF 125-284.  
 RC TISSUE=Brain;  
 RA Martin S.C., Farb D.H.;  
 RT "Molecular identification of a rat nervous system cytoplasmic  
 RT sulfotransferase, NST."  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: May catalyze the sulfate conjugation of many drugs,  
 CC xenobiotic compounds, hormones, and neurotransmitters. Displays  
 CC activity towards L-triiodothyronine, thyroxine, estrone, p-  
 CC nitrophenol, 2-naphthylamine, and 2-naphtol. May have a role in the  
 CC sulfation of drugs and neurotransmitters in the CNS (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasm.  
 CC -!- TISSUE SPECIFICITY: Expressed in the brain, not detected in the  
 CC liver, kidney, spleen, heart, small intestine or testis.  
 CC -!- DEVELOPMENTAL STAGE: Expressed at low levels in brains of 1-day  
 CC old animals but increase to adult levels from 7-day old animals  
 CC and remain at that level in adults.  
 CC -!- SIMILARITY: Belongs to the sulfotransferase family.

```
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AF188699; AAF61198.1; -; mRNA.
DR EMBL: AF176343; AAK64596.1; -; mRNA.
DR HSSP: F50224; ICDJ.
DR RGD: 69292; Sult4al.
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransferase.
KW Lipid metabolism; Steroid metabolism; Transferase.
FT CHAIN 1 284 Sulfotransferase 4A1.
FT /FTID=PRO_0000085169.
FT NP_BIND 246 254 PAPS (Potential).
SQ SEQUENCE 284 AA; 33054 MW; FCAE940F7219B6FC CRC64;

Alignment Scores:
Pred. No.: 5,91e-98 Length: 284
Score: 1506.00 Matches: 278
Percent Similarity: 99.3% Conservative: 4
Best Local Similarity: 97.9% Mismatches: 2
Query Match: 34.2% Indels: 0
DB: 1 Gaps: 0

US-10-768-158-1 (1-241) x ST4A1_RAT (1-284)
QY 21 ATGGCGGAGAGCGAGCGCGGAGACCCCGGAGTTCGAGGCAAGTACTTC 80
DB 1 MetAlaGluSerGluAlaGluThrProGlyThrProGlyGluPheGluSerIysThrPhe 20
QY 81 GAGTTCCATGGCGTGGCGTGGCGTTCCTGCGCGGGAAGATGGAGGAGATGCCCAAC 140
DB 21 GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluAspIleAlaAsp 40
QY 141 TTCGCGGTGGCGCCAGCGAGTGGATCGTACCTACCCCAAGTCGGCGACCAAGTGTG 200
DB 41 PheProValArgProSerAspValThrIleValThrIleValThrIleValThrSerLeu 60
QY 201 CTGACGAGGAGTGTACTTGTGTAGCGAGCGGCGTGCACCCCGATGAGATCGGCTTGATG 260
DB 61 LeuGlnGluValValThrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
QY 261 AACATCGACGACGCTCCGGTCTCTGAGTACCCACACGCGGCTCGGACATCATCAAG 320
DB 81 AsnIleAspGluGlnLeuProValLeuGluThrProGlnProGlyLeuAspIleIleLys 100
QY 321 GAATGACCTCTCCCGCTCATCAAGAGCACCTGCGCTACCGCTTCTGCCCTCTCAC 380
DB 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120
QY 381 CTCCACAATGAGACTCCAAGTCACTATATGCTCGCAACCCCAAGATCTGTGTGTG 440
DB 121 LeuHisAsnGlyAspSerLysValIleThrMetAlaArgAsnProLysAspLeuValVal 140
QY 441 TCTTATTATCAGTTCCACCGTCTCTGCGGACCATGAGCTACCGAGGACCTTTCAGAA 500
DB 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160
QY 501 TTCTGCGGAGGTTTATGAATGATGAAGCTGGCTACCGCTCTGTTTGAGCAGCTGAG 560
DB 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180
QY 561 GAGTTCTGGGAGCACCGATGACATCGCAACGTCGTTTTCTCAAGTATGAGACATGAT 620
DB 181 GluPheTrpGluHisArgMetAspAlaAsnValLeuPheLeuLysTyrGluAspMetHis 200
QY 621 CGGACCTGGTGACGATGGTGAGCAGTGGCCAGATTCTGTGGGGTCTCTGTGTGACAA 680
DB 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220
QY 681 GCCACGCTGGAGCCCTGACGAGGACGTCGCCACCAAGCTGGTGGACGCTGTGCAACCT 740
DB 221 AlaGlnLeuGluSerLeuIleGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240
```

```
QY 741 GAGGCCCTGCCGTCGGCGCGGAGAGCTGGCGCTGTGGAGGACATCTTACCCGCTCC 800
DB 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuThrLysAspIlePheThrValSer 260
QY 801 ATGAATGAGAAGTTGACTTGGTGTATATAACAGAGATGGGAAAGTGTGACCTCAGTTT 860
DB 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280
QY 861 GACTTTTATTTA 872
DB 281 AspPheLysLeu 284

RESULT 6
Q3TXYS MOUSE
ID Q3TXYS5 MOUSE PRELIMINARY; PRT; 284 AA.
AC Q3TXYS5
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Visual cortex cDNA, RIKEN full-length enriched library,
DE clone:K530039L09 product:sulfotransferase family 4A, member 1, full
DE insert sequence.
GN Name=Sult4al;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Visual cortex;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44 (1999).
RN [2]
NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Visual cortex;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama K., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Humnietz L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Mattiuzzi S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mortagui-Tabar S., Mulder N., Nakano N., Nakauchi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugiyara K., Sulcane R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S.C., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
```







QY 261 AACATCGAGCAGCAGCTCCCGCTCTGGAGTACCCACAGCCGGCGCTCGACATCATCAAG 320  
 Db |||||||  
 81 AaNilfAaSpGluLeuLeuProValLeuGluTyProGlnProGlyLeuAaSpfLeileIlys 100  
 QY 321 GAACTGACCTCTCCCGCTCATCAAGAGCAGCAGCTCCCGCTCTGCGCTCTGCGCTCTGAC 380  
 Db |||||||  
 101 GluLeuThrSerProArgLeuLeuIleYsSerHisLeuProTyArgPheLeuProSerAsp 120  
 QY 381 CTCCAATGGAGATCCAAGTCTATATATGGTCTGCAACCCCAAGATCTGTGTGTG 440  
 Db |||||||  
 121 LeuHisAaenGlyAaSpSerLysValIleTyMetAlaAaGAsnProLysAaSpLeuValVal 140  
 QY 441 TCTTATTATCAGTTCACCGCTCTCGGGACCATGAGTACCGAGGACACCTTTCAAGAA 500  
 Db |||||||  
 141 SerTyTyGlnPheHisArgSerLeuArgThrMetSerTyArgGlyThrPheGlnGlu 160  
 QY 501 TTCTCCCGGAGGTTTATGAATGATGAAGCTGGGCTACGGCTCTGTTGTGAGCAGCTGCAG 560  
 Db |||||||  
 161 PheCysAaArgPheMetAaSnAspLysLeuGlyTyArgLysSerTrpPheGluHisValGln 180  
 QY 561 GAGTTCCTGGAGCACCGCATGCGAAGCTGCTTTTCTCAAGTATGAAGACATGCAT 620  
 Db |||||||  
 181 GluPheTrpGluHisArgMetAspAlaAaenValLeuPheLeuLysTyGluAspMethis 200  
 QY 621 CGGACCTGTCGACATGGTGAGCAGCTGGCCAGATTCCTGGGGGTCTCTGTGACACAG 680  
 Db |||||||  
 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220  
 QY 681 GCCACGCTGGAAGCCCTGACGAGCAGCTGCCACAGCTGGTGACGCTGTGCAACGCT 740  
 Db |||||||  
 221 AlaGlnLeuGluSerLeuLeuGluHisCysHisGlnLeuValAspGlnCysAsnAla 240  
 QY 741 GAGCGCTTCGCGTGGCGGGGGAAGAGTGGGCTGTGGAAGGACATCTTCACCGTCTCC 800  
 Db |||||||  
 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAaSpIlePheThrValSer 260  
 QY 801 ATGATGAGAGTTTGACTGCTGGTGATTAACAGAGATGGGAAAGTGTGACCTTCAGTTT 860  
 Db |||||||  
 261 MetAaenGluLysPheAaSpLeuValTyLysGlnLysMetGlyLysCysAaSpLeuThrPhe 280  
 QY 861 GACTTTTATTTA 872  
 Db |||||||  
 281 AspPheTyLeu 284  
 RESULT 7  
 Q8BQV8 MOUSE  
 ID Q8BQV8 MOUSE PRELIMINARY; PRT; 284 AA.  
 AC Q8BQV8\_1  
 DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.  
 DT 07-FEB-2006, sequence version 1.  
 DT 07-FEB-2006, entry version 20.  
 DE Adult male corpora quadrigemina cDNA, RIKEN full-length enriched  
 DE library, clone:B203076P15 product:sulfotransferase family 4A, member  
 DE 1, full insert sequence.  
 GN Name=Sulf4a1;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Corpora quadrigemina; STRAIN=C57BL/6J;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Methods Enzymol. 303:19-44(1999).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
 RX PubMed=16141072; DOI=10.1126/science.1112014;  
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,  
 RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,

RA Baic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,  
 RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,  
 RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,  
 RA Bansal M., Baxtal L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,  
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,  
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,  
 RA di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,  
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,  
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,  
 RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,  
 RA Hill D., Huminick L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,  
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,  
 RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,  
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,  
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,  
 RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,  
 RA Mottaqui-Fabriz S., Mulder N., Nakano N., Nakaguchi H., Ng P.,  
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,  
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavlesi G., Pesole G.,  
 RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,  
 RA Roat B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,  
 RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,  
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,  
 RA Sperling S., Stupka E., Sugiura K., Sultana R., Takenaka Y., Taki K.,  
 RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,  
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.B., Yagi K.,  
 RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,  
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,  
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,  
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,  
 RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Nimomiya N.,  
 RA Nishio T., Okada M., Plesky C., Shibata K., Shiraki T., Suzuki S.,  
 RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,  
 RA Hayashizaki Y.;  
 RT "The transcriptional landscape of the mammalian genome.";  
 RL Science 309:1559-1563(2005).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
 RX PubMed=16141073; DOI=10.1126/science.1112009;  
 RG RIKEN Genome Exploration Research Group and Genome Science Group  
 RG (Genome Network Core Team) and the FANTOM Consortium;  
 RT "Antisense transcription in the Mammalian Transcriptome.";  
 RL Science 309:1564-1566(2005).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
 RX MEDLINE=22354683; PubMed=12468851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tonita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai S., Sasaki D., Shibata K., Shinagawa A.,

RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RL 60,770 full-length cDNAs.";  
RL Nature 420:563-573 (2002).  
RN [5]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Glissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Guatnisch S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690 (2001).  
RN [6]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RX MEDLINE=2499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayaseu N., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630 (2000).  
RN [7]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771 (2000).  
RN [8]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
RA Hori F., Imotani K., Ishii Y., Itoh M., Kaga I., Kasukawa T.,  
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,  
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NonCommercial License  
CC -----  
CC EMBL; AK046368; BAC32692.1; -; mRNA.  
CC HSSP; P50224; 1CJM.

DR Ensembl; ENSMUSG00000018865; Mus musculus.  
DR MGI; MGI:1888971; Sult4a1.  
DR GO; GO:0005737; Cytoplasm; TAS.  
DR GO; GO:0008146; F:sulfotransferase activity; IDA.  
DR GO; GO:0008146; F:sulfotransferase activity; IDA.  
DR GO; GO:0006790; P:sulfur metabolism; IDA.  
DR InterPro; IPR000863; Sulfotransferase.  
DR Pfam; PF00685; Sulfotransferase; 1.  
DR ProDom; PD001218; Sulfotransferase; 1.  
SQ SEQUENCE 284 AA; 32972 MW; 38D7B556F2B9391E CRC64;  
Alignment Scores:  
Pred. No.: 2,53e-96 Length: 284  
Score: 1483.00 Matches: 274  
Percent Similarity: 97.9% Conservative: 4  
Best Local Similarity: 96.5% Mismatches: 6  
Query Match: 33.7% Indels: 0  
DB: 2 Gaps: 0  
US-10-768-158-1 (1-2419) x Q8BQV8\_MOUSE (1-284)  
QY 21 ATGCGGAGAGCGAGCGGAGACCCCGAGCACCACCGGAGGATTCGAGAGCAAGTACTTTC 80  
DB 1 MetGlyGluSerGluAlaGluThrProGlyThrProGlyGluPheGluSerLysTrpPhe 20  
QY 81 GAGTTCATGCGCGTGGCTGCGGCTGCGGCTGCGGCGGAGAGTGGAGGATCGCCCAAC 140  
DB 21 GluPheHisGlyValArgValProPheCysArgGlyLysMetGluAspIleAlaGlu 40  
QY 141 TTCCCGGTGCGCGCCAGCGACGTGTGGATCGTCACTACCTACCCCAAGTCCGGCACCAGTTG 200  
DB 41 PheProGlyArgProSerAspValTrpIleValThrTrpProLysSerGlyThrSerLeu 60  
QY 201 CTGCGAGAGTGGTCTACTTGGTGAGCAGGCGGCTGACCCCGATGATGATCGGCTTGATG 260  
DB 61 LeuGlnGluValValTrpLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80  
QY 261 AACATCGAGCAGCAGTCCCGTCTCGGTGAGTACCCAGCCGCGGCTGGACATCATCAAG 320  
DB 81 AsnIleAspGluGlnLeuAlaValLeuGluTrpProGlnProGlyLeuAspIleLeuLys 100  
QY 321 GAATGACCTCTCCCGCTCATCAAGAGCCACCTGCGCTTCTCCCTCTGAC 380  
DB 101 GluLeuThrSerProArgLeuLysSerHisLeuProTrpArgPheLeuProSerAsp 120  
QY 381 CTCACAAATGAGACTCCAAAGTCACTATATGCTGCGCAACCCCAAGGATCTGGTGTG 440  
DB 121 LeuHisAsnGlyAspSerLysValIleTrpMetAlaArgAsnProLysAspLeuValVal 140  
QY 441 TCTTATTATCATGTTCCACCGCTCTCTCGGACCATGAGTACCGAGGACCTTTCAAGAA 500  
DB 141 SerTrpTrpGlnPheHisArgSerLeuArgThrMetSerTrpArgGlyThrPheGlnGlu 160  
QY 501 TTCTCGCGAGGTTTATGAATGATAAGTGGGCTACCGCTCTCGTGTGGACCACTGCGAG 560  
DB 161 PheCysArgArgPheMetAsnAspLysLeuGlyTrpGlySerTrpPheGluHisValGln 180  
QY 561 GAGTTCGGGAGACCGCATCGACTCGAAGCTGCTTTTCTCAAGTATGAGAGACATGAT 620  
DB 181 GluPheTrpGluHisArgMetAspAlaAsnValLeuPheLeuLysTrpGluAspMethis 200  
QY 621 CGGGACCTGGTGACGATGGTGGAGCAGCTGCCAGATTCCTGGGGGTCTCTGTGACAAG 680  
DB 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220  
QY 681 GCCCAGCTGGGAAGCCCTGACCGAGCCTGCCACCGAGTGGTGGAGCCAGTCTGCAACGCT 740  
DB 221 AlaGlnLeuGluSerLeuIleGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240  
QY 741 GAGGCCCTGCGGCGCGGAGAGAGTGTGGCTGTGGAGAGGACATCTTACCGCTCTCC 800  
DB 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 260

```
QY 801 ATGAATGAGAACTTTGACTTGTGTATAAACAGAGATGGGAAAGTGTCACTCAGTTT 860
|||||
Db 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280
|||||
QY 861 GACCTTTTATTTA 872
|||||
Db 281 AspPheTyrLeu 284
|||||

RESULT 8
ID Q68EV3_XENLA PRELIMINARY; PRT; 284 AA.
AC Q68EV3;
DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE MGC84327 protein.
GN Name=MGC84327;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
[2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RA Klein S., Gerhard D.S.; to the EMBL/GenBank/DBJ databases.
RL Submitted (AUG-2004)
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC080097; AA080097.1; --; mRNA.
DR GO; GO:0008146; F:sulfoltransferase activity; IEA.
DR InterPro; IPR000863; Sulfoltransferase.
DR Pfam; PF00685; Sulfoltransferase 1; 1.
DR ProDom; PD001218; Sulfoltransferase; 1.
SQ SEQUENCE 284 AA; 33211 MW; 9344EE54A62729D3 CRC64;

Alignment Scores:
```

```
Pred. No.: 2.28e-89 Length: 284
Score: 1385.00 Matches: 253
Percent Similarity: 95.4% Conservative: 18
Best Local Similarity: 89.1% Mismatches: 13
Query Match: 31.5% Indels: 0
DB: 2 Gaps: 0

US-10-768-158-1 (1-2419) x O68EV3_XENLA (1-284)

QY 21 ATGCGGAGAGCGAGCGGAGAGACCCAGCACCCCGGGGAGTTTCGAGAGCAAGTACTTC 80
|||||
Db 1 MetAlaGluSerGluAlaGluThrProSerThrProCysGluPheGluSerLysTyrPhe 20
|||||
QY 81 GAGTTCCATGCGTGGCGCTCCGCCCTTTCGCCGGGGAAGATGGAGAGATCCCAAC 140
|||||
Db 21 GluTyrAsnGlyIleArgLeuProPheCysArgGlyLysMetGluGluValSerAsp 40
|||||
QY 141 TTCCCGGTGGCGCCAGCGAGCTGTGGATCTCACTACCCACCCAGTCCGGGACCAAGCTTG 200
|||||
Db 41 PheProValArgGluAsnAspIlePheIleValThrTyrProLysSerGlyThrSerLeu 60
|||||
QY 201 CTGACGAGGAGTGGTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATG 260
|||||
Db 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
|||||
QY 261 AACATCGACGAGCTCCCGTCTCTGAGTACCCACAGCGGGCGCTCGACATCATCAAG 320
|||||
Db 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleLys 100
|||||
QY 321 GAATGTACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTTCGCCCTCTCAC 380
|||||
Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120
|||||
QY 381 CTCACCAATGGAGACTCCAAGTCTATATGGCTCGCAACCCCAAGATCTGTGGTG 440
|||||
Db 121 LeuHisAsnGlyAsnSerLysValIleTyrMetAlaArgAsnProLysAspLeuValVal 140
|||||
QY 441 TCTTATTATAGTTCCACCGCTCTCTCGGACCATAGCTACCGAGGACCTTTCACAGAA 500
|||||
Db 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160
|||||
QY 501 TTCTGCGGAGGTTTATGAATGATAGCTGGGTACCGCTCCTGGTTTGACACGCTGCAG 560
|||||
Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTyrPheAspHisValGln 180
|||||
QY 561 GAGTTCTGGGAGCACCGCATCGATCGAACCTGCTTTTCTCAAGTATGAAGACATGCAT 620
|||||
Db 181 GluPheTyrAspHisArgLeuAspSerAsnValLeuPheLeuLysTyrGluAspMethis 200
|||||
QY 621 CGGACCTGTGTGACGATGGTGGAGCAGCTGGCCAGATTCTCTGGGGGTCTCTGTGACAAG 680
|||||
Db 201 LysAspLeuGlyThrMetValGluGlnLeuValArgPheLeuGlyValSerTyrAspLys 220
|||||
QY 681 GCCCAGCTGGGAAGCCCTGACCGGACACTGCCACACGCTGGTGGGACCACTGTGCAACGCT 740
|||||
Db 221 AlaGlnLeuGluSerThrIleGluHisCysHisLeuLeuIleAspHisCysCysAsnAla 240
|||||
QY 741 GAGGCCCTGCGCGTGGGCGGGGAGAGATTGGGTGTGGAAGGACATCTTCAACCGCTCTCC 800
|||||
Db 241 GluAlaLeuProIleGlyArgGlyArgValGlyLeuTyrLysAspIlePheThrValSer 260
|||||
QY 801 ATGAATGAGAGTTTGAATTGGTGTATAAACAGAGATGGGAAAGTGTGACCTCAGTTT 860
|||||
Db 261 MetAsnGluLysPheAspGlnValTyrLysGlnArgMetGlyLysLeuAspLeuAlaPhe 280
|||||
QY 861 GACCTTTTATTTA 872
|||||
Db 281 GluPheAsnLeu 284
|||||

RESULT 9
Q52KW3_XENLA PRELIMINARY; PRT; 284 AA.
ID Q52KW3_XENLA
AC Q52KW3;
```



```

RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jallion O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costat C., Bernot A.,
RA Nicaud C., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva S., Salanoubat M., Levy M., Boudet N., Castellino S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Bienmont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Crüaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RL the early vertebrate proto-karyotype.";
RN Nature 431:946-957(2004).
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CAAG01014526; CAF96774.1; -; Genomic DNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
SQ SEQUENCE 284 AA; 32894 MW; 414AB27E7D8DF9A6 CRC64;

Alignment Scores:
Pred. No.: 2,94e-84 Length: 284
Score: 1313.00 Matches: 243
Percent Similarity: 93.0% Conservative: 21
Best Local Similarity: 85.6% Mismatches: 20
Query Match: 29.8% Indels: 0
DB: 2 Gaps: 0

US-10-768-158-1 (1-2419) x Q4SRF8_TETNG (1-284)
Qy 21 ATGGCGAGAGCGCGGCGAGACCCCGAGCACCCCGGGGAGTTCCGAGCAAGTACTTC 80
Db 1 MetalaalaaArgGluuAaAspThrProSerThrProIleGluPheGluSerIysTyrPhe 20
Qy 81 GAGTTCATGCGTGGCGCTCCGCGCTTCCTGCGCGGGAAGATGGAGGAGATCGCCAA 140
Db 21 GluPheAspGlyValArgLeuProPheCysArgGlyLysMetGluGluIleAlaGly 40
Qy 141 TTCGCGTGGCGCGGCGAGCGTGGATCGTACCTACCTACCCAGTCCGCGCACGCTTG 200
Db 41 PheSerLeuArgSerAspIlePheValThrTyrProLysSerGlyThrSerLeu 60
Qy 201 CTGCAAGAGGTGGTCTACTTGTGTAGCCAGGCGCTGACCCCGATGAGATCGGTTGATG 260
Db 61 LeuGlnGluIleValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
Qy 261 AACATCGACGACGCTCCGCGTCTGAGTACCCACACGCGGCGCTCGACATCATCAAG 320
Db 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleGln 100
Qy 321 GAACGTGACCTCTCCCGCTCATCAAGAGCCACCTGCGCTACCGCTTCTGCCCTCTGAC 380
Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrGlnPheLeuProThrAla 120
Qy 381 CTCACAAATGAGATCCCAAGTCACTATATGCTCGCAACCCCAAGATCTGGTGGTG 440
Db 121 LeuHisAsnGlyGluAlaLysValIleTyrMetAlaArgAsnProLysAspLeuValVal 140
Qy 441 TCTTATTATAGTTCCACCGCTCTCTGCGGACCATAGACTACCGAGGCACCTTTCAAGAA 500

```

```

Db 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160
Qy 501 TTCTCCCGGAGGTTTATGAATGATAAGTGGCTACGGCTCCTGGTTTGGACACGCTGCAG 560
Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180
Qy 561 GAGTCTGGGAGCCCGCATCGACCTGCAACGTCGCTTTTCTCAAGTATGAAGACATGCAT 620
Db 181 GluPheTrpGluHisArgMetAsnSerAsnValLeuPheLeuLysTyrGluAspMetTyr 200
Qy 621 CGGGACCTGGTGCAGATGGTGCAGACCTGCGCAGATTCCTGGGGTCTCTGTGCACAAG 680
Db 201 LysAspLeuGlyThrPheValGluGluuAlaArgPheLeuGlyValSerCysAspLys 220
Qy 681 GCCCAGCTGGAAAGCCCTGACGCGACCTGCACACGAGCTGGTGGACGACCTGCACACCT 740
Db 221 AlaGlnLeuGluGlyLeuValGluSerCysAsnGlnLeuIleGluGlnCysSerAsnSer 240
Qy 741 GAGCCCTGCGCGGCGGAGAGTTCGGGCTGTGGAGAGACATCTTCACCGCTCTCC 800
Db 241 GluAlaLeuSerValCysArgGlyArgValGlyLeuTrpLysAspValPheThrValSer 260
Qy 801 ATGAATGAGAAGTTGACTTGGTCTGTATAAACAGAGATGGGAAGTGTGACCTCAGCTT 860
Db 261 MetAsnAspLysPheAspThrIleTyrArgGlnLysMetCysLysSerAspLeuThrPhe 280
Qy 861 GACTTTTATTATTA 872
Db 281 AspPheGlyLeu 284

RESULT 11
Q5RCA8_PONPY PRELIMINARY; PRT; 218 AA.
ID Q5RCA8_PONPY PRELIMINARY; PRT; 218 AA.
AC Q5RCA8
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 21-DEC-2004, sequence version 1.
DE Hypothetical protein DKFZp459I0535.
GN Name=DKFZp459I0535;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Cortex;
RG The German cDNA Consortium;
RA Koshner K., Beyer A., Mewes H.W., Weill B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CR858370; CAH90599.1; -; mRNA.
DR SMR; Q5RCA8; 35-218.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Hypothetical protein.
SQ SEQUENCE 218 AA; 25328 MW; 87B9496A82B7B878 CRC64;

Alignment Scores:
Pred. No.: 2,86e-71 Length: 218
Score: 1129.50 Matches: 215
Percent Similarity: 81.7% Conservative: 0
Best Local Similarity: 81.7% Mismatches: 3
Query Match: 25.7% Indels: 45
DB: 2 Gaps: 1

```

```
US-10-768-158-1 (1-2419) x Q5RCB8_PONPY (1-218)
QY 88 ATGGCGTGGCGCTGCGGCTTCTGCGCGGGAAGATGGAGGAGATGCCAACTTCCGG 147
DB 1 MetAlaCysGlyCysArgProSerAlaAlaGlyArgTrpArgSerProThrSerArg 20
QY 148 TGGCGGCCAGGAGCTGTGGATCGTCACCTACCCCAAGTCGGGACACAGCTTGTGCGAG 207
DB 21 CysGlyProAlaThrCysGlyLeuSerProThrProSerPro----- 34
QY 208 AGGTGGTCTACTTGTGTAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCG 267
DB 34 ----- 34
QY 268 ACGAGCAGCTCCCGTCTCGGATGATCCACAGCCGGCGCTGGACATCATCAAGAACTGA 327
DB 35 -----Glu-LeuT 37
QY 328 CCTCTCCCGGCTCATCAAGAGCCACCTGCGCTACCGCTTTCTGCGCTCTGACCTCCACA 387
DB 37 hrSerProArgLeuIleIysSerHisLeuProTyrArgPheProProSerAspLeuHisA 57
QY 388 ATGGAGACTCAAGGTCTATATGCTGCTCGCAACCCCAAGGATCTGGTGTCTTATT 447
DB 57 snGlyAspSerIysValIleTyrMetAlaArgAsnProIysAspLeuValSerTyrT 77
QY 448 ATCAGTTCCACCGCTCTCTGCGGACCATGAGTACCGAGGACCTTTTCAAGAACTTGGCC 507
DB 77 yrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGluPheCysA 97
QY 508 GGAGTTTATGATGATGATGAGTGGGCTACGCTCTGCTTGTGAGCAGCTGAGGATGCT 567
DB 97 rGArgPheMetAsnAspIysLeuGlyTyrGlySerTrpPheGluHisValGlnGluPheT 117
QY 568 GGGACACCGGATGACTCGAACGCTGCTTTTCTCAAGTATGAGACATGATCGGGACC 627
DB 117 rpGluHisArgMetAspSerAsnValIlePheLeuLeuIysTyrGluAspMetHisArgAspL 137
QY 628 TGGTGACGATGGTGAGCAGCTGGCCAGATCTCTGGGGGTGTCTGTGACAGGCCAGC 687
DB 137 euValThrMetValGlnGlnLeuAlaArgPheLeuGlyValSerCysAspIysAlaGlnL 157
QY 688 TGGAGCCCTACCGAGCAGCTGCCACAGCTGGTGGAGCAGTGTGCAACCTGAGGCC 747
DB 157 euGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAlaGluAlaL 177
QY 748 TGCCCGTGGCGGGAAGAGTTGGCTGTGGAAGGACATCTTCAACGCTCTCCATGAATG 807
DB 177 euProValGlyArgGlyArgValGlyLeuTyrLysAspIlePheThrValSerMetAsnG 197
QY 808 AGAAGTTTGACTGGTGATTAACAGAGATGGGAAAGTGTGACCTCACGTTTGACTTTT 867
DB 197 luLysPheAspLeuValTyrIysGlnIysMetGlyLysCysAspLeuThrPheAspPheT 217
QY 868 ATTTA 872
DB 217 yrLeu 218
RESULT 12
Q9RLS5_MOUSE PRELIMINARY; PRT; 298 AA.
AC Q9RLS5;
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 18.
DE Arylsulfotransferase STIA4.
GN Name=Sultial;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
```

```
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=129SV;
RA Honma W., Nagata K., Yamazoe Y.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AB029487; BAA82321.1; -; Genomic_DNA.
DR PIR; S28183; S28183.
DR HSSP; P50224; 1CJM.
DR SMR; Q9RLS5; 12-298.
DR Ensembl; ENSMUSG00000030711; Mus musculus.
DR MGI; MGI:102896; Sultal.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 298 AA; 34717 MW; E5100E784A263668 CRC64;
Alignment Scores:
Pred. No.: 3,34e-26 Length: 298
Score: 495.00 Matches: 104
Percent Similarity: 53.9% Conservative: 54
Best Local Similarity: 35.5% Mismatches: 117
Query Match: 11.2% Indels: 18
DB: 2 Gaps: 3
US-10-768-158-1 (1-2419) x Q9RLS5_MOUSE (1-298)
QY 36 GCCGAGACCCCGACACCCCGGGGAGTTCGAGAGCAAGTACTTCGAGTTCATGGCGTG 95
DB 2 AlaGlnAsnProSerAsnMetGluProLeuArgIysProLeuValProValIysGlyLe 21
QY 96 CGGTGCGCGCTCTCTGCGCGGGAAGATGGAGGAGATCGCAACTTCCCGTGGCGGCC 155
DB 22 ProLeuLeIyryrPheAlaGluThrMetGluGlnLeuGlnAsnPheThrAlaTrpPro 41
QY 156 AGCGACGTGTGGATCGTCACCTACCCCAAGTCCCGGACCAAGCTTGTGCGAGAGTGGTC 215
DB 42 AspAspValIleIleSerThrTyrProIysSerGlyThrAsnTrpMetSerGluIleMet 61
QY 216 TACTTGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGACGAGCAG 275
DB 62 AspMetIleTyrGlnGlyGlyLysLeuAspIysCysGlyArgAlaProValIysAlaArg 81
QY 276 CTCCCGGTCTCGGAGTACCCA-----CAGCGGCGCTGGACATCATCAAGAA 323
DB 82 IleProPheLeuGluPheSerCysProGlyValProProGlyLeuGluThrLeuIysGlu 101
QY 324 CTGACCTCTCTCCCGCTCATCAAGAGCCACCTGCGCTTACCGCTTCTGCGCTCTGACCTC 383
DB 102 ThrProAlaProArgIleIleLysThrHisLeuProLeuSerLeuLeuProGlnSerLeu 121
QY 384 CACAATGGAGACTCAAGGTCTATCTATGCTGCTGCAACCCCAAGGATCTGCTGCTGCT 443
DB 122 LeuAspGlnIysIleLysValIleTyrValAlaArgAsnAlaLysAspValValSer 141
QY 444 TATTATCATGTTCCACCGCTCTCTCGGACCATGAGTACCGAGGCGACCTTTCAAGAAATTC 503
DB 142 TyrTyrAsnPheTyrLysMetAlaLysLeuHisProAspProGlyThrTrpGluSerPhe 161
QY 504 TGGCGGAGTTTATGAATGATAGCTGGGCTTACCGCTCTGCTGTTTGAACACGTCGAGAG 563
DB 162 LeuGluAsnPheMetAspGlyLysValSerTyrGlySerTrpTyrGlnHisValIysGlu 181
QY 564 TTCTGGGAGACCGGATGAGTCCGAACGTGCTTTTCTCAAGTATGAAGACATGATCGG 623
DB 182 TrpTrpGluLeuArgArgThrHisProValLeuTyrLeuPheTyrGluAspMetIysGlu 201
QY 624 GACCTGTGACGATGGTGGAGCAGCTGGCCAGATTCCTGGGGGTGCTCCTGTGACAAAGGCC 683
```

```
Db 202 AenProLysArgGluIleLysLysIleLeuGluPheLeuGlyArgSerLeuProGluGlu 221
Qy 684 CAGCTGGAAGCCCTGACGAGCAGCTGC-----CACCAGCTGGTGGACAGTGCTGC 734
Db 222 ThrValAspLeuIleValHisHisThrPheLysLysMetLysGluAsnProMetAla 241
Qy 735 AACGCTGAGGCCCTGCCC-----GTGGGCCGG 761
Db 242 AenTyThrThrIleProThrGluValMetAspHisThrIleTyThrProPheMetArgLys 261
Qy 762 GGAAGAGTGGCTGGAAGGACATCTTCCCGCTCCATGAATGAGAAGTTTTCACCTG 821
Db 262 GlyThrIleGlyAspTrpLysAsnThrPheThrValAlaGlnSerGluHisPheAspAla 281
Qy 822 GTGTATAACAGAGATGGGAAGTGTGACCTCAGCTTT 860
Db 282 HisTyAlaLysLeuMetThrGlyCysAspPheThrPhe 294

RESULT 13
ID ST1A1_MOUSE STANDARD; PRT; 291 AA.
AC P52840;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-1996, sequence version 1.
DE 07-FEB-2006, entry version 43.
DE sulfotransferase 1A1 (EC 2.8.2.1) (Aryl sulfotransferase) (Phenol
DE sulfotransferase) (Sulfokinase) (Phenol/aryl sulfotransferase) (mStp1)
DE (ST1A4).
GN Name=St1a1; Synonyms=St1a1, Stp, Stp1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [mRNA].
RC STRAIN=C57BL/6 J CBA; TISSUE=liver;
RX MEDLINE=93144346; PubMed=8424956; DOI=10.1016/0167-4781(93)90073-M;
RA Kong A.-N.T., Ma M., Tao D., Yang L.;
RT "Molecular cloning of cDNA encoding the phenol/aryl form of
RT sulfotransferase (mStp1) from mouse liver.";
RL Biochim. Biophys. Acta 1171:315-318(1993).
CC -!- FUNCTION: Catalyzes O-sulfation of phenols, N-O-sulfation of
CC minoxidil and tyrosine esters (By similarity).
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + a phenol =
CC adenosine 3',5'-bisphosphate + an aryl sulfate.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC ENBL; L02331; -, NOT_ANNOTATED_CDS; mRNA.
DR HSSP; P50224; 1CUM.
DR SMR; P52840; 5-291.
DR Ensembl; ENSMUSG00000030711; Mus musculus.
DR MGI; MGI:102896; St1a1.
DR GO; GO:0000166; F:nucleotide binding; IDA.
DR GO; GO:0008146; F:sulfotransferase activity; IDA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR Lipid metabolism; Steroid metabolism; Transferase.
FT CHAIN 1 291 Sulfotransferase 1A1.
FT /FTID=PRO_0000085130.
FT PAPS (By similarity).
FT NP_BIND 44 49 PAPS (By similarity).
FT NP_BIND 126 134 PAPS (By similarity).
FT NP_BIND 189 225 PAPS (By similarity).
FT NP_BIND 253 255 PAPS (By similarity).
FT ACT_SITE 104 104 Proton acceptor (By similarity).
FT SEQUENCE 291 AA; 33974 MW; ADC07187DFAS1D96 CRC64;
SQ
```

```
Alignment Scores:
Pred. No.: 8.84e-26 Length: 291
Score: 489.00 Matches: 101
Percent Similarity: 56.0% Conservative: 53
Best Local Similarity: 36.7% Mismatches: 103
Query Match: 11.1% Indels: 18
DB: 1 Gaps: 3

US-10-768-158-1 (1-2419) x ST1A1_MOUSE (1-291)
Qy 90 GGCCTGGCGCTGCCCGCTCTGCGCGGGAAGATGGAGGATCGCCAACTTCCCGGTG 149
Db 13 GlyIleProLeuIleLysTyThrPheAlaGluThrMetGluGlnLeuGluAsnPheThrAla 32
Qy 150 CGGCCCGCAGCGAGTGTGGATCGTCACCTACCCCAAGTCCGGCACCAGCTGTGCGAGG 209
Db 33 TrpProAspAspValLeuIleSerThrTyThrProLysSerGlyThrAsnTrpMetSerGlu 52
Qy 210 GTGTCTACTTGTGAGCCAGCGCGTACCCCGATGAGATCGGCTTGTATGAACATCGAC 269
Db 53 IleMetAspMetIleTyThrGlnGlyLysLeuAspLysCysGlyArgAlaProValTyThr 72
Qy 270 GAGCAGCTCCCGGCTCGGAGTACCCA-----CAGCCGGCGCTGACATCATC 317
Db 73 AlaArgIleProPheLeuGluPheSerCysProGlyValProProGlyLeuGluThrLeu 92
Qy 318 AAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCGCTTACCGCTTTTCTGCCCTCT 377
Db 93 LysGluThrProAlaProArgIleIleLysThrHisLeuProLeuSerLeuLeuProGln 112
Qy 378 GACCTCCACATGAGACTCCAAAGTCAATCATATATGCTCCCAACCCCAAGATCTGTG 437
Db 113 SerLeuLeuAspGlnLysIleLysValIleTyThrValAlaArgAsnAlaLysAspValVal 132
Qy 438 GTGTCTTATTATCAGTTCCACCGCTCTCTGCGGACCATGACTACCGAGGACCTTTCAA 497
Db 133 ValSerTyThrAsnPheTyThrLysMetAlaLysLeuHisProAspProGlyThrTrpGlu 152
Qy 498 GAATTTCTCGCGAGGTTTATGAATGATAAGCTGGGCTACGGCTCTCTGTTTGGACACGTG 557
Db 153 SerPheLeuGluAsnPheMetAspGlyLysValSerTyThrGlySerTyThrGlnHisVal 172
Qy 558 CAGAGTTCTGGGAGCACCAGCAGTCCGACCTGCTGTTTCTCAAGTATGAAGACATG 617
Db 173 LysGluTrpTrpGluLeuArgArgThrHisProValLeuTyThrLeuPheTyThrGluAspMet 192
Qy 618 CATCGGACCTGGTGACGATGGTGAGCAGCTGCCAGATTCCTGGGGGTCTGCTGTGAC 677
Db 193 LysGluAsnProLysArgGluIleLysLysIleLeuGluPheLeuGlyArgSerLeuPro 212
Qy 678 AAGCCCGCAGCTGGAAGCCCTGACGAGCAGCTGC-----CACCAGCTGGTGGACCCAG 728
Db 213 GluGluThrValAspLeuIleValHisHisThrSerPheLysLysMetLysGluAsnPro 232
Qy 729 TGCTGCAACGCTGAGGCCCTGCGCC-----GTG 755
Db 233 MetAlaAsnTyThrThrIleProThrGluValMetAspHisThrIleTyThrProPheMet 252
Qy 756 GGCCGGGGAAGAGTTGGCTGTGGGAAGGACATCTTCCACCGTCTCCATGAATGAGAAGTTT 815
Db 253 ArgLysGlyThrIleGlyAspTrpLysAsnThrPheThrValAlaGlnSerGluHisPhe 272
Qy 816 GACTTGTGTATAAACAAGAGATGGGAAGTGTGACCTCAGCTTT 860
Db 273 AspAlaHisTyThrAlaLysLeuMetThrThrGlyCysAspPheThrPhe 287

RESULT 14
ST1C1_MOUSE
ID ST1C1_MOUSE STANDARD; PRT; 296 AA.
AC Q9D935; O8R210;
DT 13-APR-2004, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-2001, sequence version 1.
```



07-FEB-2006, entry version 22.  
 DT Sulfotransferase 1C1 (EC 2.8.2.-).  
 DE Name=Sulf1c1;  
 GN Mus musculus (Mouse).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OX Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [mRNA].  
 RC TISSUE=Kidney;  
 RA Sugimura K.;  
 RT "cDNA cloning of a novel mouse kidney sulfotransferase associated with  
 DT DPT-induced polycystic kidney";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC STRAIN=C57BL/6J; TISSUE=Pancreas;  
 RX PubMed=16141072; DOI=10.1126/science.1112014;  
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,  
 RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,  
 RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,  
 RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,  
 RA Ambesi-Impombato A., Apweiler R., Attaluri R.N., Bailey T.L.,  
 RA Banaal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,  
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,  
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,  
 RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,  
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,  
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,  
 RA Guatinchich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,  
 RA Hill D., Humnick L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,  
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,  
 RA Kizano H., Kollas S., Krishnan S.P., Kruger A., Kummerfeld S.K.,  
 RA Kurochkin I.V., Larau L.F., Lazarevic D., Lipovich L., Liu J.,  
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,  
 RA Mateuza H., Matuzawa S., Miki H., Mignone F., Miyake S., Morris K.,  
 RA Mortagui-Tabar S., Mulder N., Nakano N., Nakauchi H., Ng P.,  
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,  
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,  
 RA Petrovsky N., Piazza S., Reid J., Reid J.F., Ring B.Z., Ringwald M.,  
 RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,  
 RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,  
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,  
 RA Sperling S., Stupka E., Sugita K., Sultana R., Takenaka Y., Taki K.,  
 RA Tannoja K., Tan S.L., Tang S.L., Taylor M.S., Tegner J., Teichmann S.A.,  
 RA Ueda H.R., van Nimwegen E., Varardo R., Wei C.L., Yagi K.,  
 RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,  
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,  
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,  
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,  
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,  
 RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,  
 RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,  
 RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,  
 RA Hayashizaki Y.;  
 RT "The transcriptional landscape of the mammalian genome";  
 RL Science 309:1559-1563 (2005).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC TISSUE=Kidney;  
 RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueffing T.B., Toshikiyuki S., Carninci P., Prange C.,  
 RA Rana S.A., Loguercio N.A., Peters G.J., Abramson R.D., Mullany P.H.,  
 RA Bosak S.J., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -!- FUNCTION: Catalyzes the sulfate conjugation of many drugs,  
 CC xenobiotic compounds, hormones, and neurotransmitters (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).  
 CC -!- SIMILARITY: Belongs to the sulfotransferase family.  
 CC  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 CC EMBL: AY005469; AAC00823.1; -; mRNA.  
 CC EMBL: AK007384; BAB25002.1; -; mRNA.  
 CC EMBL: BC022665; AAH22665.1; -; mRNA.  
 CC HSSP: P50224; 1CUM.  
 CC SMR: Q9D939; 12-296.  
 CC Ensembl: ENSMUSG00000023122; Mus musculus.  
 CC MGI: MGI:1916333; Sult1c1.  
 CC InterPro: IPR000863; Sulfotransferase.  
 CC Pfam: PF00685; Sulfotransferase; 1.  
 CC PRODOM: PD001218; Sulfotransferase; 1.  
 CC TRANSFAM: Sulfotransferase.  
 FT CHAIN 1 296 Sulfotransferase 1C1.  
 FT NP\_BIND 49 54 /FTID=PRO\_0000085133.  
 FT NP\_BIND 131 139 PAPS (By similarity).  
 FT NP\_BIND 194 230 PAPS (By similarity).  
 FT NP\_BIND 258 260 PAPS (By similarity).  
 FT ACT\_SITE 109 109 Proton acceptor (By similarity).  
 FT CONFLICT 95 95 D -> V (in Ref. 3).  
 SQ SEQUENCE 296 AA; 34953 MW; 50B8648F2A2BEA57 CRC64;  
 Alignment Scores:  
 Pred. No.: 3.87e-25 Length: 296  
 Score: 480.00 Matches: 103  
 Percent Similarity: 55.4% Conservative: 57  
 Best Local Similarity: 35.6% Mismatches: 107  
 Query Match: 10.9% Indels: 22  
 DB: 1 Gaps: 5  
 US-10-768-158-1 (1-2419) x ST1C1\_MOUSE (1-296)  
 QY 66 GAGAGCAGTACTTCGAGTTCATGCGGTGGCGGTGGCGGCTTCGCGCGGGAAGATG 125  
 Db 10 GlnThrLysLeuLysGluValAlaGlyIleProLeuGlnAlaProThrValAspAsnTrp 29  
 QY 126 GAGGAGATCGCAACTTCCCGTGGCGGCCAGCAGCAGTGTGGATCGTCACCTACCCCAAG 185  
 Db 30 ArgGlnIleGlnThrPheGluAlaLysProAspAspLeuLeuIleCysThrTyProLys 49  
 QY 186 TCCGCGCACCATCTGTCGAGGAGGTGTCTACTTGTGGACGCGGCGCTGACCCCGAT 245  
 Db 50 SerGlyThrThrTrpIleGlnGluValAspMetIleGluGlnAsnGlyAspValGlu 69  
 QY 246 GAGATCGGCTTGATGACATCGACGACGCTCCCGGTCTCGTGGAGTAC-----CCA 296  
 Db 70 LysCysArgArgThrIleIleGlnHisArgHisProPheIleGluTrpAlaArgProPro 89  
 QY 297 CAGCCG---GGCCTGGACATCATCAAGAACTGACCTCTCCCGCCTCATCAAGAGCAC 353  
 Db 90 GlnProSerGlyValAspLysAlaAsnGluMetProAlaProArgIleLeuArgThrHis 109  
 QY 354 CTGCGCTACCGCTTCTGCGCTCTGACCTCCACATGAGACTCCAGGTCTATCTATG 413  
 Db 110 LeuProThrGlnLeuLeuProProSerPheThrTrpAsnAsnLysPheLeuTyVal 129







DE sulfotransferase 1A3) (Sulfotransferase, monoamine-preferring) (M-BST)  
DE (Thermolabile phenol sulfotransferase) (TL-EST) (Placental estrogen  
DE sulfotransferase) (Catecholamine-sulfating phenol sulfotransferase  
DE (HAST3).  
GN Name=SULT1A3; Synonyms=STM;  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).  
RC TISSUE=Brain;  
RX MEDLINE=93371391; PubMed=8363392;  
RA Zhu X., Veronese M.E., Bernard C.C., Sansom L.N., McManus M.E.;  
RT "Identification of two human brain aryl sulfotransferase cDNAs.";  
RL Biochem. Biophys. Res. Commun. 195:120-127(1993).  
RN [2]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).  
RC TISSUE=Placenta;  
RX MEDLINE=94244843; PubMed=8187949; DOI=10.1016/0303-7207(94)90159-7;  
RA Bernier F., Lopez-Solache I., Labrie F., Luu-The V.;  
RT "Cloning and expression of cDNA encoding human placental estrogen  
RT sulfotransferase.";  
RL Mol. Cell. Endocrinol. 99:R11-R15(1994).  
RN [3]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).  
RC MEDLINE=95100963; PubMed=7802665;  
RA Dooley T.P., Probst P., Munroe P.B., Moie S.E., Liu X., Doggett N.A.;  
RT "Genomic organization and DNA sequence of the human catecholamine-  
RT sulfating phenol sulfotransferase gene (STM).";  
RL Biochem. Biophys. Res. Commun. 205:1325-1332(1994).  
RN [4]  
RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 84-101.  
RC TISSUE=Liver;  
RX MEDLINE=94161723; PubMed=8117269;  
RA Wood T.C., Aksoy I.A., Aksoy S., Weinshilboum R.M.;  
RT "Human liver thermolabile phenol sulfotransferase: cDNA cloning,  
RT expression and characterization.";  
RL Biochem. Biophys. Res. Commun. 198:1119-1127(1994).  
RN [5]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).  
RC TISSUE=Platelet;  
RX MEDLINE=95209696; PubMed=7695637;  
RA Aksoy I.A., Weinshilboum R.M.;  
RT "Human thermolabile phenol sulfotransferase gene (STM): molecular  
RT cloning and structural characterization.";  
RL Biochem. Biophys. Res. Commun. 208:786-795(1995).  
RN [6]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).  
RC TISSUE=Blood;  
RX MEDLINE=95209704; PubMed=7695643;  
RA Jones A.L., Hagen M., Coughtrie M.W.H., Roberts R.C., Glatt H.;  
RT "Human platelet phenol sulfotransferases: cDNA cloning, stable  
RT expression in V79 cells and identification of a novel allelic variant  
RT of the phenol-sulfating form.";  
RL Biochem. Biophys. Res. Commun. 208:855-862(1995).  
RN [7]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).  
RC TISSUE=Leukocyte;  
RX MEDLINE=95050600; PubMed=7961757;  
RA Bernier F., Leblanc G., Labrie F., Luu-The V.;  
RT "Structure of human estrogen and aryl sulfotransferase gene. Two mRNA  
RT species issued from a single gene.";  
RL J. Biol. Chem. 269:28200-28205(1994).  
RN [8]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).  
RC TISSUE=Liver;  
RA Gaidigk A., Grant D.M.;  
RT Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
RN [9]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).  
RX PubMed=14702039; DOI=10.1038/ng1285;

RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
RA Sekine M., Ohashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,  
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,  
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,  
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,  
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,  
RA Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y., Ishida S.,  
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,  
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y.,  
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,  
RA Yushino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,  
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,  
RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,  
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,  
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y.,  
RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,  
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,  
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,  
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;  
RT "Complete sequencing and characterization of 21,243 full-length human  
RT cDNAs.";  
RL Nat. Genet. 36:40-45(2004).  
RN [10]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).  
RC TISSUE=Lung, and Pancreas;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne S.J.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [11]  
RP NUCLEOTIDE SEQUENCE OF 139-198.  
RC TISSUE=Lymphocyte;  
RX MEDLINE=95130098; PubMed=7829089;  
RA Aksoy I.A., Callen D.F., Apostolou S., Her C., Weinshilboum R.M.;  
RT "Thermolabile phenol sulfotransferase gene (STM): localization to  
RT human chromosome 16p11.2.";  
RL Genomics 23:275-277(1994).  
RN [12]  
RP CHARACTERIZATION  
RX MEDLINE=94379981; PubMed=8093002;  
RA Veronese M.E., Burgess W., Zhu X., McManus M.E.;  
RT "Functional characterization of two human sulfotransferase cDNAs that  
RT encode monoamine- and phenol-sulfating forms of phenol  
RT sulfotransferase: substrate kinetics, thermal-stability and  
RT inhibitor-sensitivity studies.";  
RL Biochem. J. 302:497-502(1994).  
RN [13]

Db	77	ValArgValProPheLeuGluValAsnAspProGlyGluProSerGlyLeuGluThrLeu	96
Qy	318	AAGAACTGACCTCTCCCGCCTCATCAAGAGCCACCTGCCCTACCGCTTCTCGCCCTCT	377
Db	97	LysAspThrProProArgLeuIleLysSerHisLeuProLeuAlaLeuLeuProGln	116
Qy	378	GACCTCCACATGAGACTCAAGGTCACTATATATGGCTCGCAACCCCAAGCATCTGGTG	437
Db	117	ThrLeuLeuAspGlnLysValLysValValThrValAlaAlaArgAsnProLysAspValAla	136
Qy	438	GTGCTCTATTATCAGTTCCACCGCTCTCGCGGACCATGAGCTACCGAGCCACCTTTCAA	497
Db	137	ValSerTyTyHisPheHisArgMetGluLysAlaHisProGluProGlyThrThrAsp	156
Qy	498	GAATTCGCGGAGTTTATGAATGATTAAGCTGGCTACGGCTCTCTGGTTTGAGACATG	557
Db	157	SerPheLeuGluLysPheMetAlaGlyGluValSerTyGlySerTyGlnHisVal	176
Qy	558	CAGGAGTTCGGGAGACCGCATGGACTCGAAGCTGCTTTTCTCAAGTATCAAGACATG	617
Db	177	GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyLeuPheTyGluAspMet	196
Qy	618	CATCGGACCTGGTGCAGATGTCGAGCAGCTGGCCAGTTCCTGGGGGTGCTGTGTGAC	677
Db	197	LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro	216
Qy	678	AAGCCCGACGTGGAAGCCCTGACGGAGCAGCTGC-----	710
Db	217	GluGluThrMetAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro	236
Qy	711	-----CACGAGCTGGTGACCACTGCTGCAACGCTGAGGCC	746
Db	237	MetThrAsnTyThrThrValProGlnGluLeuMetAspHisSerIleSer-----	253
Qy	747	CTGCCC---GTGGCGCGGGGAAGAGTTGGCGTGTGGGAAGGACATCTTCACCTCTCCATG	803
Db	254	---ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGln	272
Qy	804	AATCAGAAAGTTTGACTTGGTCTATAACAGAAGATGGGAAAGTGTGACCTCACGTTT	860
Db	273	AsnGluArgPheAspAlaAspTyAlaGluLysMetAlaGlyCysSerLeuSerPhe	291
RESULT 17			
STK1	RAT	STANDARD;	PRT; 296 AA.
ID	STK1	RAT	
AC	Q9WU8;		
DT	01-DEC-2000,	integrated into UniProtKB/Swiss-Prot.	
DT	01-NOV-1999,	sequence version 1.	
DT	21-FEB-2006,	entry version 32.	
DT	Sulfotransferase K1 (EC 2.8.2.-)	(rSULT1C2).	
GN	Name=Sult1c1; Synonyms=Sult1c2;		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;		
OC	Muridea; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE [MRNA].		
RC	TISSUE=Kidney;		
EX	MEDLINE=20323235; PubMed=10872834; DOI=10.1006/bbrc.2000.2744;		
RA	Xiangrong L., Joehnk C., Hartmann D., Schestag F., Kroemer W.,		
RA	Gieselmann V.;		
RT	"Enzymatic properties, tissue-specific expression, and lysosomal		
RT	location of two highly homologous rat SULT1C2 sulfotransferases.";		
RL	Biochem. Biophys. Res. Commun. 272:242-250(2000).		
CC	-!- FUNCTION: Catalyzes the sulfate conjugation of many drugs,		
CC	xenobiotic compounds, hormones, and neurotransmitters.		
CC	-!- SUBCELLULAR LOCATION: Lysosome.		
CC	-!- TISSUE SPECIFICITY: Highly expressed in kidney and at lower levels		
CC	in stomach and liver. More specifically found in the epithelia of		
CC	proximal tubules of the kidney, of the bile duct, of the gastric		
CC	mucosa, and in hepatocytes.		
CC	-!- SIMILARITY: Belongs to the sulfotransferase family.		

```

Alignment Scores:
Pred. No.: 1.43e-24 Length: 295
Score: 472.00 Matches: 100
Percent Similarity: 54.5% Conservative: 52
Best Local Similarity: 35.8% Mismatches: 101
Query Match: 10.7% Indels: 26
DB: 1 Gaps: 4

US-10-768-158-1 (1-2419) x ST1A3_HUMAN (1-295)

Qy 90 GCGCTGGCGCTCGCCCTTCCTCGCGGGGAAGATGGAGGATCGCAACTTCCCGGTG 149
      ||||| ||||| :::: :::: :::: :::: ::::
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36

Qy 150 CGGCCCGACGAGTGTGGATGTCACCTACCCCAAGTCGGCACCGACTTGTCTGCAGGAG 209
      ||||| :::: ||||| :::: ||||| :::: ||||| ::::
Db 37 ArgProAspAspLeuLeuIleAsnThrTyrProLysSerGlyThrThrTyrValSerGln 56

Qy 210 GTGGTCTACTTGGTGAGCCGGCCCTCACCCCGATGAGATCGCTTGTAGAACATCGAC 269
      ::::: ::::: ||||| ||||| ::::: |||||
Db 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysAsnArgAlaProIleTyr 76

Qy 270 GAGCAGCTCCCGGTCTCTGGAGTATCCACAGCCGCGGCTTGGAGCATCATC 317
      ::::: ::::: ||||| ||||| ::::: |||||

```

```
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AJ238391; CAB41460.1; -; mRNA.
DR PIR; JC7282; JC7282.
DR HSP; F50224; ICDM.
DR RGD; 621064; Sult1c2.
DR Ensembl; ENSRNOG00000031833; Rattus norvegicus.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Lysosome; Transferase.
FT CHAIN 1 296 Sulfotransferase K1.
FT NP_BIND 49 54 /FTID-PRO_0000085135.
FT NP_BIND 131 139 PAPS (By similarity).
FT NP_BIND 194 230 PAPS (By similarity).
FT NP_BIND 258 260 PAPS (By similarity).
FT ACT_SITE 109 109 Proton acceptor (By similarity).
SQ SEQUENCE 296 AA; 34699 MW; 4A82FAD740EBABF5 CRC64;

Alignment Scores:
Pred. No.: 3 24e-24 Length: 296
Score: 467.00 Matches: 100
Percent Similarity: 55.4% Conservative: 58
Best Local Similarity: 35.1% Mismatches: 105
Query Match: 10.6% Indels: 22
DB: 5 Gaps: 5

US-10-768-158-1 (1-2419) x STK1_RAT (1-296)
QY 66 GAGAGCAAGTACTTCGAGCTTCATGGCGTGGCGCTCGCGCCCTTCGCGGGGGAAGATG 125
D 10 GlnThrLysLeuLysGluValAlaGlyIleProLeuGlnAlaProThrValAspAsnTrp 29
QY 126 GAGGAGATCGCAACTTCGCGTGGCGGCCGAGCAGCTGTGGATCGTCACCTACCCCAAG 185
D 30 SerGlnIleGlnThrPheLysAlaLysProAspLeuLeuIleCysThrTyrProLys 49
QY 186 TCCGSCACACGCTTCGACGAGGTGTCTACTTGTGGTGGAGCCAGCGGCGTACCCCGAT 245
D 50 SerGlyThrThrTrpIleGlnGluIleValAspMetIleGlnGlnAsnGlyAspValGlu 69
QY 246 GAGATCGGCTTGATGAACATGACGAGCAGCTCCCGCTCGGTCCGGAGTAC-----CCA 296
D 70 LysCysGlnArgThrIleIleGlnHisArgHisProPheIleGluTrpAlaArgProPro 89
QY 297 CAGCCG---GGCTGGACATCATCAAGGAATGACCTCTCCCGCTCATCAAGAGCCAC 353
D 90 GlnProSerGlyValAspLysAlaAsnAlaMetProAlaProArgIleLeuArgThrHis 109
QY 354 CTGCGCTACCGCTTCTCGCCCTCTGACCTCCACATGGAGACTCCCAAGTCTATATG 413
D 110 LeuProThrGlnLeuLeuProSerPheThrThrAsnAsnGlyPheLeuThrVal 129
QY 414 GCTCGCAACCCCAAGGATCTGGTGTCTTATTATTCAGTCTCCACCGCTCTCTCGGACC 473
D 130 AlaArgAsnAlaLysAspCysMetValSerTyrTyrHisPheTyrArgMetSerGlnVal 149
QY 474 ATGAGCTACCGAGGACCTTTCAAGATTTCTCGCGAGGTTTATGAATGATAGCTGGGC 533
D 150 LeuProAspProGlyThrThrAsnGluTyrPheGluThrPheIleAsnGlyLysValSer 169
QY 534 TACGCTCTCTGGTTGACACGTGACGAGTCTCGGAGTCTCGGAGCACCAGCTCGAACCTG 593
D 170 TrpGlySerTrpPheAspHisValLysGlyTrpTrpGluIleArgAspArgTyrGlnIle 189
QY 594 CTTTTCCTCAAGTATGAACATGATCGGACCTCGGACCTGGTGACGATGGAGGAGCTGCC 653
D 190 LeuPheLeuPheTyrGluAspValLysArgAspProLysArgGluIleGlnLysValMet 209
QY 654 AGATTCTCTGGGGGTGTCC-----TGTGACAGGCCCGCCAGCTGGAGCCCTG 698
```

```
Db 210 GlnPheMetGlyLysAsnLeuAspGluValValAspLysIleValLeuGluThrSer 229
QY 699 ACGGAGCAGCTGCCACCGAGCTGGTGCACAGCTGCAGCTGAGCGCTGCC----- 752
D 230 PheGluLysMetLysGlu-----AsnProMetThrAsnArgSerThrValProLysSer 247
QY 753 -----GTGGGCGGGGAGAGAGATTGGCTGTGAAGAC 785
D 248 ValLeuAspGlnSerIleSerProPheMetArgLysGlyThrValGlyAspTrpLysAsn 267
QY 786 ATCTTCACCGCTCCATCAATGACAGTTGACCTGGTATTAACAGACAGATGGAAAG 845
D 268 HisPheThrValAlaGlnAsnAspArgPheAspGluIleTyrLysGlnLysMetGlyGly 287
QY 846 TGTGACCTTCAGCTTT 860
D 288 ThrSerLeuAsnPhe 292

RESULT 18
Q3ZAV3 RAT PRELIMINARY; PRT; 296 AA.
AC Q3ZAV3;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DE 07-FEB-2006, entry version 3.
DE Sulfotransferase family, cytosolic, 1C, member 2.
GN Name=Sult1c2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RG NIH MGC Project;
RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC103636; AA103637.1; -; mRNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
KW Transferase.
SQ SEQUENCE 296 AA; 34699 MW; 4A82FAD740EBABF5 CRC64;

Alignment Scores:
```



```

Db 50 SerGlyThrThrTrpIleGlnIleValAenMetIleGlnIleAenGlyAspValGlu 69
QY 246 GAGATCGGCTTGAATCAACATGACGAGCAGCTCCCGGCTCGAGTAC-----CCA 296
Db 70 LysCysGlnArgThrIleIleGlnHisArgHisProValIleGluTrpAlaArgPro 89
QY 297 GAGCGG----GGCTGGACATCATCAAGAACTGACCTCTCCCGGCTCATCAAGACCCAC 353
Db 90 GlnProSerGlyValAspLysAlaAenAlaMetProAlaProArgIleLeuArgThrHis 109
QY 354 CTGCGCTACCGCTTCTGCGCTCTGACCTCCACATGAGAGCTCAAGTGCATCATATG 413
Db 110 LeuProGlnLeuLeuProSerPheThrPheThrAsnAsnCysLysIleuTrpVal 129
QY 414 GCTCGCAACCCCAAGGATCTGGTGTCTTATTATCATGTTCCACCGCTCTCTCGGACC 473
Db 130 AlaArgAenAlaLysAspCysMetValSerTyThrHisPheTyArgMetCysGlnVal 149
QY 474 ATGACTACCGAGGACCTTCAAGATTCCTCCGGAGGTTTATGAATGATAGCTGGGC 533
Db 150 LeuProAenProGlyThrTrpAsnGluTyPheGluThrPheIleAenGlyLysValSer 169
QY 534 TACGCTCTCTGTTGACAGCTGCAGAGTTCTCGGAGCACCGCATGAGCTCGAACCTG 593
Db 170 TrpGlySerTrpPheAspHisValLysGlyTrpTrpGluIleArgAspArgTyTrpGlnIle 189
QY 594 CTTTTCCTCAAGTATCAAGACATGATCGGACCTCGGTGACGATGCTGGAGCAGCTGCC 653
Db 190 LeuPheLeuPheTyGluAspMetLysArgAspProLysArgGluIleGlnLysValMet 209
QY 654 AGATTCTCGGGGTGTC-----TGTGACAGGCCCGACCTGGAAGCCCTG 698
Db 210 GlnPheMetGlyLysAenAlaAspGluGluValValAspLysIleValLeuGluThrSer 229
QY 699 ACGGAGCACTGCCACCACTGCTGTCAGCAGCTGCTCAACGCTGAGGCCCTGCC----- 752
Db 230 PheGluLysMetLysAsp-----AsnProLeuThrAenPheSerThrIleProLysThr 247
QY 753 -----GTGGCGCGGGAAGAGTTGGGCTGTGGAGGAC 785
Db 248 IleMetAspGlnSerIleSerProPheMetArgLysGlyIleValGlyAspTrpLysAen 267
QY 786 ATCTTCACGCTCTCCATGAATGAGAGTTTGATCTGGTATATAACAGAAATGGGAAG 845
Db 268 HisPheThrValAlaGlnAenGluArgPheAspGluIleTyGluGlnLysMetAspGly 287
QY 846 TGTGACCTCACGTTT 860
Db 288 ThrSerLeuAenPhe 292

RESULT 20
Q569D0_RAT
ID Q569D0_RAT PRELIMINARY; PRT; 296 AA.
AC Q569D0;
DT 10-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 10-MAY-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Sulfotransferase family, cytosolic, 1C, member 1 (Predicted).
GN Name=Sulfic1_predicted;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1];
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=2388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

```

```

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalsky U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2];
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RG NIH MGC Project;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC092564; AAH92564.1; -; mRNA.
DR SMR; Q569D0; 12-296.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1; 1.
KW Transferase.
SQ SEQUENCE 296 AA; 34875 MW; BB0D65640B8332A3 CRC64;

Alignment Scores:
Pred. No.: 4,49e-24 Length: 296
Score: 465.00 Matches: 101
Percent Similarity: 55.9% Conservative: 59
Best Local Similarity: 35.3% Mismatches: 102
Query Match: 10.6% Indels: 24
DB: Gaps: 5

US-10-768-158-1 (1-2419) x Q569D0_RAT (1-296)
QY 66 GAGACGAAGTACTTCGAGTTCCTCCGCTCGGCTCGCGCTTCTGCGCGGGAAGATG 125
Db 10 GlnThrLysLeuLysGluValAlaGlyIleProLeuMetAspProThrValAsnAsnTrp 29
QY 126 GAGGAGATCGCCAACTTCCCGTCCGCCGAGCAGCTGTGGATCGTACCTACCCCAAG 185
Db 30 SerGlnIleGlnThrPheLysAlaLysProAspAspLeuLeuIleCysThrTyProLys 49
QY 186 TCCGCGACCACTGCTGAGGAGTGTCTACTTGTGGTGGAGCGGCTGACCCCGCAT 245
Db 50 SerGlyThrThrTrpIleGlnGluIleValAsnMetIleGluGlnAsnGlyAspValGlu 69
QY 246 GAGATCGGCTTGAATGAACATGACGAGCAGCTCCCGGCTCTGGAGTAC-----CCA 296
Db 70 LysCysGlnArgThrIleIleGlnHisArgHisProValIleGluTrpAlaArgPro 89
QY 297 GAGCGG---GGCTGGACATCATCAAGAACTGACCTCTCCCGGCTCATCAAGACCCAC 353
Db 90 GlnProSerGlyValAspLysAlaAenAlaMetProAlaProArgIleLeuArgThrHis 109
QY 354 CTGCGCTACCGCTTCTGCGCTCTGACCTCCACATGAGAGCTCAAGTGCATCATATG 413
Db 110 LeuProGlnLeuLeuProSerPheThrPheThrAsnAsnCysLysIleuTrpVal 129
QY 414 GCTCGCAACCCCAAGGATCTGGTGTCTTATTATCATGTTCCACCGCTCTCTCGGACC 473
Db 130 AlaArgAenAlaLysAspCysMetValSerTyThrHisPheTyArgMetCysGlnVal 149
QY 474 ATGACTACCGAGGACCTTCAAGATTCCTCCGGAGGTTTATGAATGATAGCTGGGC 533

```



Db	150	LeuProAsnProGlyThrTrpAsnGluTyrPheGluThrPheIleAsnGlyLysValSer	169		
Qy	534	TACGGCTCTCTGTTTGGACACGTCAGGAGTCTGGGAGCACCGCACTGGACTCGAACGCTG	593		
Db	170	TrpGlySerTrpPheAspHisValLysGlyTrpIrpGluMetArgAspArgTyrGlnIle	189		
Qy	594	CTTTTCTCAAGTATCAAGACATGCATCGGACCTGGTGTGACGTGTGGAGCAGCTGGCC	653		
Db	190	LeuPheLeuPheTyrGluAspMetLysArgAspProLysArgGluIleGlnLysValMet	209		
Qy	654	AGATCTCTGGGGGTGTC-----TGTGACAAGGCCACGCTGGAGCCCTG	698		
Db	210	GlnPheMetGlyLysAsnLeuAspGluValValAspLysIleValLeuGluThrSer	229		
Qy	699	ACGGAGCACTGCCACCAG-----CTGGTG	722		
Db	230	PheGluLysMetLysAspAsnProLeuThrAsnPheSerThrIleProLysThrIleMet	249		
Qy	723	GACCACTGTGCAACGCTGAGCCCTCGCCCGGGGGAAGAGTTGGGCTGTGGAAG	782		
Db	250	AspGln-----SerIleSerLeuPheMetArgLysGlyIleValGlyAspTrpLys	266		
Qy	783	GACATCTTCCCGTCTCCATGAATGAGAAGTTTGACTTGTGTATAAACAAGATGGGA	842		
Db	267	AsnHisPheThrValAlaGlnAsnGluArgPheAspGluIleTyrGluGlnLysLeuAsp	286		
Qy	843	AAGGTGTCACCTCAGCTTT	860		
Db	287	GlyThrSerLeuAsnPhe	292		
RESULT 21					
Q2TAB3	HUMAN				
ID	Q2TAB3_HUMAN	PRELIMINARY;	PRT; 295 AA.		
AC	Q2TAB3;				
DT	24-JAN-2006,	integrated into UniProtKB/TrEMBL.			
DT	24-JAN-2006,	sequence version 1.			
DT	07-FEB-2006,	entry version 2.			
DE	Sulfotransferase family, cytosolic, 1A, phenol-prefering, member 3.				
DE	Name=SU1T1A3;				
GN	Homo sapiens (Human).				
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;				
OC	Homo.				
OX	NCBI_TaxID=9606;				
RP	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RC	TSSUE=Brain;				
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;				
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,				
RA	Klauber R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,				
RA	Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,				
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,				
RA	Boeck S.A., McSwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,				
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,				
RA	Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Fahey J., Heaton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,				
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,				
RA	Schmarch A., Schein J.E., Jones S.J.M., Marra M.A.;				
RT	"Generation and initial analysis of more than 15,000 full-length human				
RT	and mouse cDNA sequences."				
PL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).				
RL	[2]				
RP	NUCLEOTIDE SEQUENCE.				
RC	TISSUE=Brain;				
RG	NIH MGC Project;				



```

Db 273 AsnGluArgPheAspAlaAspTyrAlaGlnLysMetAlaGlyCysSerLeuSerPhe 291
STBI1 CHICK
ID STBI1_CHICK STANDARD; PRT; 296 AA.
AC Q8UG30;
DT 04-JAN-2005, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-2002, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-)
DE (Sulfotransferase 1B1).
GN Name=SULT1B1; Synonyms=SULT1B;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND FUNCTION.
RC TISSUE=Lymphoid;
RX PubMed=15234270; DOI=10.1016/j.abb.2004.05.008;
RA Wilson L.A., Reynolds G.E., Darras V.M., Coughtrie M.W.H.;
RT "cDNA cloning, functional expression, and characterization of chicken
sulfotransferases belonging to the SULT1B and SULT1C families.";
RL Arch. Biochem. Biophys. 428:64-72(2004).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of many hormones,
neurotransmitters, drugs and xenobiotic compounds. Sulfonation
increases the water solubility of most compounds, and therefore
their renal excretion, but it can also result in bioactivation to
form active metabolites. Sulfates thyroid hormones including 3,3'-
diiodothyronine, catechols such as 4-nitrocatechol and 4-
isopropylcatechol and phenols such as 2-bromophenol.
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; AJ494980; CAD41949.1; -; mRNA.
DR HSP; P49888; IHV3.
DR DR
DR Ensembl; ENSGALG00000011812; Gallus gallus.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR Lipid metabolism; Steroid metabolism; Transferase.
FT CHAIN 1 296 Sulfotransferase family cytosolic 1B
member 1.
FT NP_BIND 48 53 /FTID=PRO_0000085164.
FT NP_BIND 131 139 PAPS (By similarity).
FT NP_BIND 194 230 PAPS (By similarity).
FT ACT_SITE 109 109 PAPS (By similarity).
FT BINDING 259 259 Proton acceptor (By similarity).
FT BINDING 259 259 PAPS (By similarity).
SQ SEQUENCE 296 AA; 34074 MW; 487AB790416995F5 CRC64;

Alignment Scores:
Pred. No.: 5.28e-24 Length: 296
Score: 464.00 Matches: 102
Percent Similarity: 52.4% Conservative: 53
Best Local Similarity: 34.5% Mismatches: 108
Query Match: 10.5% Indels: 33
DB: 1 Gaps: 5

US-10-768-158-1 (1-2419) x STBI1_CHICK (1-296)
Qy 59 GGAGTTCGAGCAAGTACTTCGAGTTCAGGCTGGCGTGGCGTGGCGCCCTT----- 109
Db 2 GlyThrValAspAlaTyrLeuArgGlnProTrpSerThrValHisAlaIleProMetVal 21
Qy 110 -----CTGCCCGGGAGAGATGGAGGATCGCCACTCCCGTGGCGGCCAGCAGCT 163
Db 22 SerAlaPheAlaGln-AsnTrpGluArgValAspAsnPheGlnSerArgProAspIle 41

```

```

Qy 164 GTGGATCGTCACCTACCCCAAGTCCGGCACCAGCTTCTCGCAGGAGGTGTCTACTTGGT 223
Db 41 eValValAlaThrPheProLysSerGlyThrThrTrpIleSerGluIleValAspMetIle 61
Qy 224 GAGCCAGGCGCTGACCCCGATGAGATCGGCTTCATGAACATCGACGAGCTCCCGGT 283
Db 61 eLeuGlnGlyGlyAspProLysLysCysLysAspAlaIleValAsnArgValProMe 81
Qy 284 CTGGAGTACCCACAGCCG-----GGCTGGACATCATCAAGAACTTCAC 328
Db 81 tLeuGluPheAlaAlaProGlyGlnMetProAlaGlyThrGlnGlnLeuGluAsnMetPr 101
Qy 329 CTCTCCCGCTCATCAAGAGCCACCTGCGCTTACCGCTTCTGCTCCCTCTGACCTCCCAA 388
Db 101 oSerProArgIleIleLysThrHisIleProAlaAspIleLeuProLysSerPheTrpAs 121
Qy 389 TGGAGACTCCAAAGTCTATATGGCTCGCACCCCAAGGATCTGCTGGTGTCTTATTA 448
Db 121 pLysSerCysLysMetIleTyrValGlyArgAsnAlaValAspValIleSerTyrTy 141
Qy 449 TCAGTTCACCGCTCTCTGCGGACCATGAGCTACCGAGGACCTTTCAGAAATCTCGCG 508
Db 141 rHisPheAspLeuMetAsnLysLeuHisProHisProGlyThrTrpAspGlnTyrLeuGl 161
Qy 509 GAGTTTATGAATGATGAAGCTGGCTACGGCTCTCTGTTTCAGCAGCTGAGGAGTCTG 568
Db 161 uAlaPheMetAlaGlyLysValAlaTyrGlySerTrpPheAspHisValArgGlyTyrTr 181
Qy 569 GGAGCACCGCATGAGCTCGAACGCTGCTTTTCTCAAGTATGACAGCATGCTCGGACCT 628
Db 181 pGluArgArgGlnGluHisProLeuLeuTyrLeuPheTyrGluAspMetLysGluAspLe 201
Qy 629 GGTGACCATGTTGAGCAGCTGCGCAGATTCCTGGGGGTGTCTGTGACAAAGGCCACGCT 688
Db 201 uArgArgGluValAlaLysValAlaGlnPheLeuGlyArgGluLeuThrGluValAlaLe 221
Qy 689 GGAAGCCCTGACGGAGCAGCTGTC----- 710
Db 221 uAspAlaIleAlaHisThrSerPheGluAlaMetArgAspAsnProSerThrAsnTy 241
Qy 711 -----CACCAGCTGGTGACCATGCTGCAACGCTGAGGCCCTGCCCC---GT 754
Db 241 rSerValValProSerGlnLeuMetAspHisGlyIleSer-----ProPheMe 257
Qy 755 GGGCCGGGGAAGAGTTCGGCTGGGAAGACATCTTCACCGCTCTCCATGAATGAGAAGTT 814
Db 257 tArgLysGlyIleThrGlyAspTrpLysAsnHisPheThrValAlaGlnSerAlaHisPh 277
Qy 815 TGACTTGGTGTATAAACAAGATGGGAAAGTGTGACCTCACGTTT 860
Db 277 eAspGlnTyrTyrAlaGlnLysMetAlaGlyThrAspLeuArgPhe 292

RESULT 23
Q58CV8_BOVIN PRELIMINARY; PRT; 298 AA.
ID Q58CV8_BOVIN AC Q58CV8;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Sulfotransferase family, cytosolic, 1C, member 2.
GN Name=SULT1C2;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Pooled;
RX MEDLINE=21180013; PubMed=11282978; DOI=10.1101/gr.170101;
RA Smith T.P.L., Grosse W.M., Freking B.A., Roberts A.J., Stone R.T.,
RA Casas E., Wray J.E., White J., Cho J., Fahrenkrug S.C., Bennett G.L.,

```

RA Heaton M.P., Laegreid W.W., Rohrer G.A., Chitko-McKown C.G.,  
 RA Perteau G., Holt I., Karamycheva S., Liang F., Quackenbush J.,  
 RA Keele J.W.;  
 RT "Sequence evaluation of four pooled-tissue normalized bovine cDNA  
 RT libraries and construction of a gene index for cattle.";  
 RL Genome Res. 11:626-630(2001).  
 RN (2)  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Poolled;  
 RA Hathay G.P., Sonstegard T.S., Van Tassel C.P., Clawson M.L.,  
 RA Heaton M.P., Keele J.W., Snelling W.M., Weidmann R.T., Smith T.P.L.;  
 RT "Sequencing and analysis of Bos taurus full-length insert cDNA  
 RT clones";  
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC (2)  
 DR EMBL; BT021839; AAX46686.1; -; mRNA.  
 DR SMR; Q58CV8; 11-262.  
 DR Ensembl; ENSBTAG00000016032; Bos taurus.  
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR InterPro; IPR000863; Sulfotransferase.  
 DR Pfam; PF00685; Sulfotransfer\_1; 1.  
 KW Transferase.  
 SQ SEQUENCE 298 AA; 34914 MW; E597E9EA1032E37E CRC64;  
 Alignment Scores:  
 Pred. No.: 6,23e-24 Length: 298  
 Score: 463.00 Matches: 98  
 Percent Similarity: 56.1% Conservatives: 58  
 Best Local Similarity: 35.3% Mismatches: 102  
 Query Match: 10.5% Indels: 20  
 DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x Q58CV8\_BOVIN (1-298)

QY 90 GCGTGGCGTCCGCG---CCCTTCTGCGCGGGAGATGGAGGATCGCCAACTTCCG 146  
 DQ 19 GlytLeuGlnProLysProThrCys---AspThrTrpAspGlnIleTrpSerPheGln 37  
 QY 147 GTGGCGCCACGACGTGTGTGATCGTACCTACCCCAAGTCCGCGACGAGTCTGTGAG 206  
 DQ 38 AlaArgProAspAspLeuLeuIleSerThrTyrProLysAlaGlyThrThrTrpGln 57  
 QY 207 GAGTGTGTCTACTTGGTGAGCGGCGCTGACCCCGATGAGATCGGCTTGATGAACATC 266  
 DQ 58 GluIleValAspLeuIleGlnAsnGlyGlyAspValAsnGlnSerGlnArgAlaProThr 77  
 QY 267 GACGAGCAGCTCCCGGTCTCGAGTACCCACGCGG-----GGCTGGACATC 314  
 DQ 78 HisGluArgPheProPheIleGluTrpIleProSerLeuGlySerGlyLeuGluGln 97  
 QY 315 ATCAAGGAACGACTCTCCCGCTCATCAAGACCGCTCCCTACCGCTTTCTGCGC 374  
 DQ 98 AlaAsnAlaMetAlaSerProArgMetLeuLysThrHisLeuProPheHisLeuLeuPro 117  
 QY 375 TCTGACCTCCCAATGGAGCTCCCAAGGTCTATATATGGTCCGCAACCCCAAGATCTG 434  
 DQ 118 ProSerPheLeuGluAsnCysLysMetIleTyrValAlaArgAsnProLysAspAsn 137  
 QY 435 GTGTGTCTTATTATCAGTCTCCCGCTCTCTCGGACCATGAGCTACCGAGGACCTTT 494  
 DQ 138 MetValSerTyrTyrHisPheHisArgMetAsnArgAsnLeuProAlaProGlyThrTrp 157  
 QY 495 CAAGAATCTCCGAGGTCTTATGAATCATAGCTGGCTACGGCTCTGTTGAGCAC 554  
 DQ 158 GluGluTyrPheGluSerPheLeuAlaGlyLysValCysTrpGlySerTrpTyrAspHis 177  
 QY 555 GTGCAGGAGTCTCTGGGACACCGCATGAGCTCGAAGCTGCTTTTCTCAAGTATGAAGAC 614  
 DQ 178 ValLysGlyTyrTrpGlnAlaLysAspGlnHisArgIleLeuTyrLeuPheTyrGluAsp 197

QY 615 ATGCATCGGACCTGGTGACGATGGTGAGCAGCTGGCCAGATTCCTGGGGGTCTCCTGT 674  
 DQ 198 MetLysGluAsnProLysHisGluIleGlnLysLeuAlaGluPheIleGlyLysSerLeu 217  
 QY 675 GACAAGGCCCGACGTGGAGCCCTGACGAGCACTGCCACGAGCTGGT-----GAC 725  
 DQ 218 AspAspLysValLeuAspLysIleValAspHisThrSerPheSerValMetLysGlnAsn 237  
 QY 726 CAGTGTGCAACGCTGAGGCCCTCGCC----- 752  
 DQ 238 ProMetAlaAsnTyrThrSerIleProAsnGluTyrMetAsnGlnLeuIleSerProPhe 257  
 QY 753 GTGGCGCGGAGAGTGGCTGTGGAAGGACATCTTCACCGTCTCCATGAATGAGAAG 812  
 DQ 258 MetArgLysGlyValIleGlyAspTrpLysAsnHisPheThrValAlaGlnAsnGluArg 277  
 QY 813 TTTGACTGTGTATAAACAGAGATGGAAAGTGTGACCTCAGCTTGTGACTTT 866  
 DQ 278 PheAspAspAspTyrArgLysAsnMetAlaAspThrThrLeuThrLeuHisPhe 295

RESULT 24  
 ST1A1\_RAT  
 ID ST1A1\_RAT STANDARD; PRT; 291 AA.  
 AC P17988;  
 DT 01-NOV-1990, integrated into UniProtKB/Swiss-Prot.  
 DT 01-NOV-1990, sequence version 55.  
 DE Sulfotransferase 1A1 (EC 2.8.2.1) (Aryl sulfotransferase) (Phenol  
 DE sulfotransferase) (PST-1) (Sulfokinase) (Aryl sulfotransferase IV)  
 DE (ASTIV) (Tyrosine-ester sulfotransferase) (Minoxidil  
 DE sulfotransferase).  
 GN Name=St1a1; Synonyms=St1a1;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridea; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [MRNA].  
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;  
 RX MEDLINE=90326537; PubMed=2374726;  
 RA Ozawa S., Nagata K., Gong D., Yamazoe Y., Kato R.;  
 RT "Nucleotide sequence of a full-length cDNA (PST-1) for aryl  
 RT sulfotransferase from rat liver.";  
 RL Nucleic Acids Res. 18:4001-4001(1990).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].  
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;  
 RX MEDLINE=94131305; PubMed=829966; DOI=10.1016/0378-1119(93)90028-2;  
 RA Khan A.S., Taylor B.R., Chung K., Echeredge J., Gonzales R.,  
 RA Ringer D.P.;  
 RT "Genomic structure of rat liver aryl sulfotransferase IV-encoding  
 RT gene.";  
 RL Gene 137:321-326(1993).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE OF 7-291.  
 RX MEDLINE=92379794; PubMed=1511441;  
 RA Yerkun T., Ethersedge J.L., Norton T.R., Carter H.A., Chung K.H.,  
 RA Birckbichler P.J., Ringer D.P.;  
 RT "Characterization of a complementary DNA for rat liver aryl  
 RT sulfotransferase IV and use in evaluating the hepatic gene transcript  
 RT levels of rats at various stages of 2-acetylaminofluorene-induced  
 RT hepatocarcinogenesis.";  
 RL Cancer Res. 52:4779-4786(1992).  
 RN [4]  
 RP CHARACTERIZATION.  
 RC TISSUE=Liver;  
 RX MEDLINE=93191709; PubMed=8447833;  
 RA Cruickshank D., Sansom L.N., Veronese M.E., Mojarrahi B.,  
 RA McManus M.E., Zhu X.;  
 RT "cDNA expression studies of rat liver aryl sulphotransferase.";  
 RL Biochem. Biophys. Res. Commun. 191:295-301(1993).



RA Mao C., Sanchez R.I., Clairmont K., Coughtrie M.W.H., Kauffman P.C.;  
RT "Cloning, Bacterial Expression and Characterization of Rat Brain  
RT Phenol Sulfotransferase SUT1A1: An Enzyme Involved in Neurosteroid  
RT and Dopamine Sulfonation";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN: Sprague-Dawley; TISSUE: Brain;  
RA Mao C.G., Sanchez R.I., Kauffman P.C.;  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; AF394793; AAK77559.1; -, mRNA.  
DR SMR; Q548D2; 4-291.  
DR GO; GO:0004062; F: aryl sulfotransferase activity; IEA.  
DR GO; GO:0016740; F: transferase activity; IEA.  
DR InterPro; IPR000863; Sulfotransferase.  
DR Pfam; PF00685; Sulfotransferase.1; 1.  
DR ProDom; PD001218; Sulfotransferase; 1.  
KW Transferase.  
SQ SEQUENCE 291 AA; 33906 MW; 9BC66C72923D9872 CRC64;  
  
Alignment Scores:  
Pred. No.: 1,4e-23 Length: 291  
Score: 458.00 Matches: 96  
Percent Similarity: 55.6% Conservative: 57  
Best Local Similarity: 34.9% Mismatches: 104  
Query Match: 10.4% Indels: 18  
DB: 2 Gaps: 4  
  
US-10-768-158-1 (1-2419) x Q548D2\_RAT (1-291)  
QY 90 GCGGTGCGGCTGCCCGCTTCTGCGCGGGAAGATGGAGAGATGCGCAACTTCCCGGTG 149  
DB 13 GlylleProLeuIleLysTy:PheAlaGluThrIleGlyProLeuGlnAsnPheThrAla 32  
QY 150 CGGCCACGAGGTGTGGATCGTCACCTACCCCAAGTCCGGACACGAGTCTCTGCGAGAG 209  
DB 33 TrpProAspLeuLeuIleSerThrTyProLysSerGlyThrTrpMetSerGlu 52  
QY 210 GTGGTCTACTTGGTGGACCGGCGCTGACCCGATGAGATGCGGCTTGCATCAACATCGAC 269  
DB 53 IleLeuAspMetIleTyGlnGlyLysLeuGluLysCysGlyArgAlaProIleTy 72  
QY 270 GAGCAGTCCCGGTCTCGAGTACCCACAGCCG-----GGCCTGGACATCATC 317  
DB 73 AlaArgValProPheLeuGluPheLysCysProGlyValProSerGlyLeuGluThrLeu 92  
QY 318 AAGAACTGACTCTCCCGCTCATCAAGACCCAGCTGCTTCTGCGCTCT 377  
DB 93 GluGluThrProAlaProArgLeuLeuLysThrHisLeuProLeuSerLeuLeuProGln 112  
QY 378 GACTCCCAATGAGACTCAAGTCTATATATGCTCGCAACCCCAAGATCTGGT 437  
DB 113 SerLeuLeuAspGlnLysValIleTyIleAlaArgAsnAlaLysAspValVal 132  
QY 438 GTGTCTTATTATCATGTTCCCGCTCTCTCGGACCATGAGCTACCGAGGACCTTTCAA 497  
DB 133 ValSerTyTyAsnPheTyAsnMetAlaLysLeuHisProAspProGlyThrTrpAsp 152  
QY 498 GAATTCGCGGAGTTTATGAATGATAAGCTGAGCTAGGCTCTGTTTGGACACGTG 557  
DB 153 SerPheLeuGluAsnPheMetAspGlyGluValSerTyGlySerTyTrpTyGlnHisVal 172  
QY 558 CAGGAGTTCTGGAGCACCGATGAGCTCGCAACGCTTTTCTCAAGATATGAGACATG 617  
DB 173 LysGluTrpTrpGluLeuArgHisThrHisProValLeuTyLeuPheTyGluAspIle 192  
QY 618 CATCGGACCTGGTGACGATGGAGCAGCTGCCAGATTCCTGGGGGTGCTCTGTGAC 677  
DB 193 LysGluAsnProLysArgGluIleLysLysIleLeuGluPheLeuGlyArgSerLeuPro 212

QY 678 AAGSCCAGCTGGAGCCCTGACGAGCACTGC-----CACCACTGGTGGACCACTGC 731  
DB 213 GluGluThrValAspSerIleValHisThrSerPheLysLysMetLysGluAsnCys 232  
QY 732 ---TGCAACGCTGAGGCCCTGCCC-----GTG 755  
DB 233 MetThrAsnTyThrThrIleProThrGluIleMetAspHisAsnValSerProPheMet 252  
QY 756 GCGCGGGAAGAGTTGGCTGTGGAGGACATCTCACCTGCTCCATGATGAGAGTTT 815  
DB 253 ArgLysGlyThrThrGlyAspTrpLysAsnThrPheThrValAlaGlnAsnGluArgPhe 272  
QY 816 GACTGTGTATATAACAGAGATGGGAAAGTGTGACCTCAGCTTT 860  
DB 273 AspAlaHisTyAlaLysThrMetThrAspCysAspPheLysPhe 287  
  
RESULT 26  
ID ST1A2 HUMAN STANDARD; PRT; 295 AA.  
AC P50226; P78393;  
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.  
DT 01-OCT-1996, sequence version 1.  
DE 07-FEB-2006, entry version 51.  
DE Sulfotransferase 1A2 (EC 2.8.2.1) (Aryl sulfotransferase 2) (Phenol  
DE sulfotransferase 2) (Phenol-sulfating phenol sulfotransferase 2) (P-  
DE PST 2) (ST1A2).  
GN Name=ST1A2; Synonyms=STP2;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN NUCLEOTIDE SEQUENCE.  
RC TISSUE=Liver;  
RA Zhu X., Veronese M.E., Iocco P., McManus M.E.;  
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Liver;  
RX MEDLINE=96065417; PubMed=7581483;  
RA Ozawa S., Nagata K., Shimada M., Ueda M., Tazuki T., Yamazoe Y.,  
RA Kato R.;  
RT "Primary structures and properties of two related forms of aryl  
RT sulfotransferases in human liver.";  
RL Pharmacogenetics 5:S135-S140(1995).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=8033246; PubMed=8033246; DOI=10.1016/0009-2797(94)90057-4;  
RA Yamazoe Y., Nagata K., Ozawa S., Kato R.;  
RT "Structural similarity and diversity of sulfotransferases.";  
RL Chem. Biol. Interact. 92:107-117(1994).  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=96298636; PubMed=8661000; DOI=10.1006/geno.1996.0216;  
RA Her C., Raftogiannis R., Weinshilboum R.M.;  
RT "Human phenol sulfotransferase STP2 gene: molecular cloning,  
RT structural characterization, and chromosomal localization.";  
RL Genomics 33:409-420(1996).  
RN [5]  
RP NUCLEOTIDE SEQUENCE.  
RX PubMed=8912648; DOI=10.1006/bbrc.1996.1628;  
RA Dooley T.P., Huang Z.;  
RT "Genomic organization and DNA sequences of two human phenol  
RT sulfotransferase genes (STP1 and STP2) on the short arm of chromosome  
RT 16.";  
RL Biochem. Biophys. Res. Commun. 228:134-140(1996).  
RN [6]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=97237043; PubMed=9119390; DOI=10.1006/geno.1996.4575;  
RA Gaedigk A., Beatty B.G., Grant D.M.;  
RT "Cloning, structural organization, and chromosomal mapping of the

```
RT human phenol sulfotransferase STP2 gene." ;
RL Genomics 40:242-246(1997).
RN [7]
RP CHARACTERIZATION.
RX MEDLINE=95196680; PubMed=7889867;
RY Yamazoe Y, Ozawa S, Nagata K., Gong D.-W., Kato R.;
RA "Characterization and expression of hepatic sulfotransferase involved
RT in the metabolism of N-substituted aryl compounds." ;
RL Environ. Health Perspect. 102:99-103(1994).
RN [8]
RP VARIANT ASN-235.
RX MEDLINE=2022641; PubMed=10762004;
RY DOI=10.1097/0000571-200003000-00008;
RA Engelke C.E., Meinel W., Boeing H., Glatt H.;
RT "Association between functional genetic polymorphisms of human
RL sulfotransferases 1A1 and 1A2." ;
RX Pharmacogenetics 10:163-169(2000).
RY
RA -!- FUNCTION: Catalyzes the sulfate conjugation of catecholamines,
CC phenolic drugs and neurotransmitters. Is also responsible for the
CC sulfation and activation of minoxidil. Mediates the metabolic
CC activation of carcinogenic N-hydroxyarylamines to DNA binding
CC products and could so participate as modulating factor of cancer
CC risk.
CC
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + a phenol =
CC adenosine 3',5'-bisphosphate + an aryl sulfate.
CC
CC -!- SUBUNIT: Homodimer (By similarity).
CC
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U28170; AAB09659.1; -; mRNA.
DR EMBL; U28169; AAB09658.1; -; mRNA.
DR EMBL; X78282; CAA55088.1; -; mRNA.
DR EMBL; U34804; AAB09758.1; -; Genomic_DNA.
DR EMBL; U72202; AAB08970.1; -; Genomic_DNA.
DR EMBL; U72196; AAB08970.1; JOINED; Genomic_DNA.
DR EMBL; U72197; AAB08970.1; JOINED; Genomic_DNA.
DR EMBL; U72198; AAB08970.1; JOINED; Genomic_DNA.
DR EMBL; U72199; AAB08970.1; JOINED; Genomic_DNA.
DR EMBL; U72200; AAB08970.1; JOINED; Genomic_DNA.
DR EMBL; U72201; AAB08970.1; JOINED; Genomic_DNA.
DR EMBL; U72203; AAB08970.1; JOINED; Genomic_DNA.
DR EMBL; U76619; AAB18753.1; -; Genomic_DNA.
DR EMBL; U33886; AAC51149.1; -; Genomic_DNA.
DR PIR; G01843; G01843.
DR PIR; JC5249; JC5249.
DR HSSP; P50224; 1CJM.
DR SMR; P50226; 8-295.
DR
DR Ensembl; ENSG00000197165; Homo sapiens.
DR HGNC; HGNC:11454; SULT1A2.
DR
DR MIM; 601292; gene.
DR
DR GO; GO:0008146; F:sulfotransferase activity; TAS.
DR GO; GO:0009309; P:amine biosynthesis; TAS.
DR GO; GO:0008663; Sulfotransferase.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1.
DR ProDom; PD00218; Sulfotransferase; 1.
DR Catecholamine metabolism; Lipid metabolism; Polymorphism;
KW Steroid metabolism; Transferase.
KW CHAIN 1 295 Sulfotransferase 1A2.
FT /FTID=PRO_0000085128.
FT NP_BIND 48 53 PAPS (By similarity).
FT NP_BIND 130 138 PAPS (By similarity).
FT NP_BIND 193 229 PAPS (By similarity).
FT NP_BIND 257 259 PAPS (By similarity).
FT ACT_SITE 108 108 Proton acceptor (By similarity).
FT VARIANT 7 7 T -> I.
FT /FTID=VAR_007426.
FT VARIANT 235 235 T -> N.
FT /FTID=VAR_007427.
FT CONFLICT 19 19 P -> L (in Ref. 2).
FT CONFLICT 290 290 S -> T (in Ref. 6).
FT SEQUENCE 295 AA; 34286 MW; FAD88E348B49BESA CRC64;
SQ
```

---

```
Alignment Scores:
Pred. No.: 1.95e-23 Length: 295
Score: 456.00 Matches: 94
Percent Similarity: 54.5% Conservative: 58
Best Local Similarity: 33.7% Mismatches: 101
Query Match: 10.4% Indels: 26
DB: 1 Gaps: 4

US-10-768-158-1 (1-2419) x ST1A2_HUMAN (1-295)

Qy 90 GGCCTGGCGGTCGCCGCTTCTGCGCGGGAAGATGGAGGATCGCAACTTCCGGTG 149
Dy ||||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 17 GlyValProLeuLeuLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36

Qy 150 CGGCCACGACGTCGTGGATCGTCACCTACCTACCCAGTCGCGCACGAGTCTGCGAGG 209
Dy ||||| ||||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 37 ArgProAspLeuLeuLeuSerThrTyrProLysSerGlyThrTyrValSerGln 56

Qy 210 GTGTCTACTTGTGGAGCCAGGCGGTGACCCCGATGAGATCGGCTTGATGAACATGAC 269
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysHisAlaProIlePhe 76

Qy 270 GAGCAGCTCCCGGTCCTGGAGTACCCACAGCG- - - - -GGCTGGACATCATC 317
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 77 MetArgValProPheLeuGluPheLysValProGlyIleProSerGlyMetGluThrLeu 96

Qy 318 AAGGAACCTGACCTCTCCCGCTCATCAAGAGCCACCTGCTACCGCTTTCTGCCCTCT 377
Dy ||||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 97 LysAsnThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116

Qy 378 GACCTCCACATGAGACTCCAGTCAATGCTATATGCTCGCAACCCCAAGATCTGTG 437
Dy ||||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspValAla 136

Qy 438 GTGCTTATTATCAGTTCACCGCTCTCTCGGACCATGAGTCTACCGAGGACCTTCAA 497
Dy ||||| ||||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValTyrProHisProGlyThrTrpGlu 156

Qy 498 GAATTCCTCGCGAGGTTTATGAATGAATAAGTACGGCTCTGTTGTGAGCAGCTG 557
Dy ||||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 157 SerPheLeuGluLysPheMetAlaGlyGluValSerTyrGlySerTyrGlnHisVal 176

Qy 558 CAGGAGTCTGGGAGCACCAGCATGCGAATCTGTTTCTCAAGATATGAAGACATG 617
Dy ||||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196

Qy 618 CATCGGACCTGGTGACGATGGTGGACAGCTGCCAGATTCTCTGGGGGTCTCTGTGAC 677
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216

Qy 678 AAGGCCACGCTGGAAGCCCTGACCGACACTGC- - - - - 710
Dy : : : : : : : : : : : : : : : : : : : : : : : :
Db 217 GluGluThrValAspLeuMetValGluHisThrSerPheLysGluMetLysLysThrPro 236

Qy 711 - - - - -CACAGCTGGTGACCGACAGTCTGCAACGCTGAGGCC 746
Dy : : : : : : : : : : : : : : : : : : : : : : : :
Db 237 MetThrAsnTyrThrThrValArgGluPheMetAspHisSerIleSer- - - - - 253

Qy 747 CTGCCC- - -GTGGCGCGGGAAGATTGGGCTGGGAAGGACATCTTCACGCTCTCCATG 803
Dy ||||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 254 - - -ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGln 272

Qy 804 AATGAGAAGTTTGACTTGGTGTATAAAACAGAAGATGGAAAGTGTGACCTCACCTTT 860
Dy ||||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 27
Q6Y0X5.PIG
ID Q6Y0X5.PIG PRELIMINARY; PRT; 295 AA.
AC Q6Y0X5;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
```











-----  
 Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 Distributed under the Creative Commons Attribution-NoDerivs License





```
Qy 558 CAGGAGTCTCGGAGACCCGATGCGACTGCGAAGCTGCTTTTCTCAAGTATGAAGACATG 617
Db 177 ArgglutTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheGluAspMet 196
Qy 618 CATCGGACCTGGTGTGACGATCGTGAGCAGCTGCCAGATTCCTCGGGGCTCTCTGTGAC 677
Db 197 LysGluAsnProLysArgGluLeuLysLysLeuGluPheLeuGlyArgSerLeuPro 216
Qy 678 AAGCCCGAGCTGGAAGCCCTGACGAGCAGCTGCCACCGACTGGTGGACCCAG----- 728
Db 217 GluGluThrValAspArgIleAla--HisCysThrSerPheLysGluMetLysLysAsn 235
Qy 729 ---TGCTGCAACGCTGAGGCCCTGCCCC----- 752
Db 236 ProMetThrAenTyrSerThrIleProGluAsnIleMetAspHisAsnValSerProPhe 255
Qy 753 GTGGCGCGGGAAGAGTGGCTGTGGAAGACATCTTCACCGTCTCATGAATGAGAAG 812
Db 256 MetArgArgGlyValAlaGlyAspTrpLysThrThrPheThrValAlaGlnHisGluTyr 275
Qy 813 TTTCAGCTTGGTGTATAAACAGAGATGGGAAGTGTGACCTCACGTTT 860
Db 276 PheGluAlaAspTyrAlaGluLysMetAlaGlyCysGluLeuThrPhe 291

RESULT 35
ST1C1_HUMAN
ID ST1C1_HUMAN STANDARD; PRT; 296 AA.
AC 000338;
DT 01-DEC-2000, integrated into UniProtKB/Swiss-Prot.
DT 01-JUL-1997, sequence version 1.
DT 07-MAR-2006, entry version 52.
DE Sulfotransferase 1C1 (EC 2.8.2.-) (SULT1C#1) (ST1C2) (humsULTC2).
GN Name=SULT1C1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Fetal liver, and Fetal spleen;
RX MEDLINE=97312707; PubMed=9169148; DOI=10.1006/geno.1997.4683;
RA Her C., Kaur G.P., Athwal R.S., Weinshilboum R.M.;
RT "Human sulfotransferase SULT1C1: cDNA cloning, tissue-specific
expression, and chromosomal localization.";
RL Genomics 41:467-470(1997).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Fetal liver;
RX MEDLINE=98297811; PubMed=9635988; DOI=10.1093/carcin/19.5.951;
RA Yoshinari K., Nagata K., Shimada M., Yamazoe Y.;
RT "Molecular characterization of ST1C1-related human sulfotransferase.";
RL Carcinogenesis 19:951-953(1998).
RN [3]
RN NUCLEOTIDE SEQUENCE, AND CHARACTERIZATION.
RC TISSUE=Fetal lung;
RX MEDLINE=99069375; PubMed=9852044; DOI=10.1074/jbc.273.51.33929;
RA Sakakibara Y., Yanagisawa K., Katafuchi J., Ringer D.P., Takami Y.,
RA Nakayama T., Suiko M., Liu M.-C.;
RT "Molecular cloning, expression, and characterization of novel human
SULT1C sulfotransferases that catalyze the sulfonation of N-hydroxy-2-
acetylaminofluorene.";
RL J. Biol. Chem. 273:33929-33935(1998).
RN [4]
RN NUCLEOTIDE SEQUENCE, AND CHARACTERIZATION.
RC TISSUE=Stomach;
RX MEDLINE=99410886; PubMed=10481272; DOI=10.1016/S1357-2725(99)00038-2;
RA Hehonah N., Zhu X., Brix L., Bolton-Grob R., Barnett A., Windmill K.,
RA McManus M.;
RT "Molecular cloning, expression, localisation and functional
characterisation of a rabbit SULT1C2 sulfotransferase.";
RL Int. J. Biochem. Cell Biol. 31:869-882(1999).
RN [5]
```

```
RP NUCLEOTIDE SEQUENCE, AND ALTERNATIVE SPLICING.
RX MEDLINE=20247255; PubMed=10783263; DOI=10.1006/geno.2000.6150;
RA Freimuth R.R., Rafkogianis R.B., Wood T.C., Moon E., Kim U.-J., Xu J.,
RA Sciliano M.J., Weinshilboum R.M.;
RT "Human sulfotransferases SULT1C1 and SULT1C2: cDNA characterization,
gene cloning, and chromosomal localization.";
RL Genomics 65:157-165(2000).
RN [6]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM SHORT).
RA Kalline N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA Koundinya M., Raphael J., Moreira D., Kelley T., Labaer J., Lin Y.,
RA Phelan M., Farmer A.;
RT "Cloning of human full-length CDSs in BD Creator(TM) system donor
vector.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN [7]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Kidney;
RX MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Maman A.C., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of many drugs,
xenobiotic compounds, hormones, and neurotransmitters. May be
involved in the activation of carcinogenic hydroxylamines. Shows
activity towards p-nitrophenol and N-hydroxy-2-acetylmino-
fluorene (N-OH-2AAP).
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Short;
CC IsoId=O000338-1; Sequence=Displayed;
CC Names=Long;
CC IsoId=O000338-2; Sequence=VSP_006303;
CC -!- TISSUE SPECIFICITY: Found in adult stomach, kidney and thyroid
gland, and in fetal kidney and liver.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; U66036; AAC51285.1; -; mRNA.
CC EMBL; AB008164; BAA28346.1; -; mRNA.
CC EMBL; AF026303; AAC00409.1; -; mRNA.
CC EMBL; AF186251; AAF72799.1; -; mRNA.
CC EMBL; AF186252; AAF72800.1; -; mRNA.
CC EMBL; AF186253; AAF72801.1; -; mRNA.
CC EMBL; AF186254; AAF72802.1; -; mRNA.
CC EMBL; AF186255; AAF72803.1; -; mRNA.
CC EMBL; AF186256; AAF72804.1; -; mRNA.
CC EMBL; AF186262; AAF72805.1; -; Genomic DNA.
CC EMBL; AF186258; AAF72805.1; JOINED; Genomic DNA.
CC EMBL; AF186260; AAF72805.1; JOINED; Genomic DNA.
CC EMBL; AF186261; AAF72805.1; JOINED; Genomic DNA.
CC EMBL; AF186262; AAF72806.1; -; Genomic DNA.
CC EMBL; AF186258; AAF72806.1; JOINED; Genomic DNA.
```





```

Qy 594 CTTTCTCAAGTATGAGACATGCATCGGACCTGGTGGACGATGGTGAGCAGCTGGCC 653
Db 196 LeuTyrLeuPheTyrGluAspMetLysLysAsnProLysHisGluIleGlnLysLeuAla 215
Qy 654 AGATTCTCGGGGTGCTCTGTGCAAGGCCAGCTGGAAGCCCTGACGGAGCACTGCCAC 713
Db 216 GluPheIleGlyLysLysLeuAspLysValLeuAspLysIleValHisTyrThrSer 235
Qy 714 CAGCTGGTG-----GACAGTGTGCAACGCTGAGGCCCTCGCC----- 752
Db 236 PheAspValMetLysGlnAsnProMetAlaAsnTyrSerSerIleProAlaGluIleMet 255
Qy 753 -----GTGGCGGGGAGAGTGGCTGTGGAAGGACATCTTC 791
Db 256 AspHisSerIleSerProPheMetArgLysGlyAlaValGlyAspTrpLysLysHisPhe 275
Qy 792 ACCGTCTCCATGAATGAGAGCTTTGACTGTGTATATAACAGAGATGGGAAAGTGTGAC 851
Db 276 ThrValIleGlnAsnGluArgPheAspGluAspTyrLysLysMetThrAspThrArg 295
Qy 852 CTCAGTTGACTTT 866
Db 296 LeuThrPheHisPhe 300

RESULT 37
Q95MF8_PIG PRELIMINARY; PRT; 295 AA.
AC Q95MF8;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 14.
DE Estrogen sulfotransferase.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA KIM J.G., Vallet J.L., Rohrer G.A., Christenson R.K.;
RX MEDLINE=22345564; PubMed=12457956; DOI=10.1016/S0739-7240(02)00172-8;
RT "Characterization of porcine uterine estrogen sulfotransferase.";
RL Domest. Anim. Endocrinol. 23:493-506(2002).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AF389855; AAK72967.1; -, mRNA.
DR HSP; P49888; 1HY3.
DR SMR; Q95MF8; 4-293.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR Transference.
SQ SEQUENCE 295 AA; 34940 MW; 7677C28A91B38167 CRC64;

Alignment Scores:
Pred. No.: 8,49e-23 Length: 295
Score: 447.00 Matches: 96
Percent Similarity: 54.3% Conservative: 56
Best Local Similarity: 34.3% Mismatches: 106
Query Match: 10.2% Indels: 22
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x Q95MF8_PIG (1-295)
Qy 66 GAGAGCAAGTACTTCGAGTTCT-----CATGGGTGGCGGTGGCGCCCTTCTGCG 113
Db 5 LysSerAlaTyrLeuAspTyrPheGlyArgIleHisGlyIleLeuLeuTyrLysLysPhe 24

```

```

Qy 114 CGCGGGAAGATGGAGGATCGCCAACTTCCGGTGGCGCCAGCAGCGTGTGGATCGTC 173
Db 25 IleGluTyrTrpAsnAspValGluThrPheGluAlaArgProAspLeuVal 44
Qy 174 ACCTACCCCAAGTCCGCGCACCGCTTCTGTCAGAGGTGTCTACTTGGTGGACCGGCGC 233
Db 45 ThrTyrProLysSerGlyThrTyrTrpValSerGluIleValTyrMetIleTyrThrGlu 64
Qy 234 GCTGACCCCATGATGATCGCTGTGATGAACATGACGAGCAGCTCCCGTCCTG----- 287
Db 65 GlyAspValGluLysCysLysGluAspThrIlePheAsnArgIleProTyrLeuGluCys 84
Qy 288 -----GAGTACCACAGCCGCGCTGGACATCATCAAGGAAGTCACTCTCCCGCGCTC 341
Db 85 ArgThrGluAsnValMetAsnGlyValLysGlnLeuLysGlnMetAlaSerProArgIle 104
Qy 342 ATCAAGAGCCACCTGCCCTACCGCTTCTGCTCCCTCTGACCTCCACATGAGACTCCCAAG 401
Db 105 ValLysSerHisLeuProProGluLeuLeuProValSerPheTrpGluLysAsnCysLys 124
Qy 402 GTCATCTATATGCTGCCAACCCCAAGGATCTGGTGTGTCTATTATCATGATTCACCGC 461
Db 125 IleIleTyrValCysArgAsnAlaLysAspValValSerTyrTyrPhePheLeu 144
Qy 462 TCTCTCGGACCATGAGTACCGGACCTTTCAAGATTTCTCGCGAGGTTTATGAAT 521
Db 145 MetValThrAlaAsnProAspProGlySerPheGlnAspPheValGluLysPheMetAsp 164
Qy 522 GATAAGCTGGCTACGGCTCTGTTTGTAGCAGCTGCAGAGTTCCTGGGAGCACCGCATG 581
Db 165 GlyGluValProTyrGlySerTyrTyrLysHisThrLysSerTrpTrpGluLysArgThr 184
Qy 582 GACTCGAACCTGCTTTTCTCAAGTATGAAGACATGATCGGACCTGGTGTGATCGTGTG 641
Db 185 AsnProGlnValLeuPheIlePheTyrGluAspMetLysGluAsnIleArgLysGluVal 204
Qy 642 GAGCAGCTGGCCAGATTCTCTGGGGGTCTCTGTGACAAGCCCGACCTGGAGCCCTGACG 701
Db 205 MetArgLeuIleGluPheLeuGlyArgLysAlaSerAspGluLeuValAspLysIle 224
Qy 702 GAGCAGCTGCCAC-----CAGCTGGTG 722
Db 225 LysHisThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrThrLeuPro 244
Qy 723 GACAGTGTGCAACGCTGAGCGCTGCGCC-----GTGGCGCGGAGAGATGGCTGTGG 779
Db 245 AspGluValMetAsnGlnLysValSerAlaPheMetArgLysGlyIleAlaGlyAspTrp 264
Qy 780 AAGGACATCTTCAACGCTCTCCATGAATGAGAAGTTTGACTTGTGTATAACAGAACATG 839
Db 265 LysAsnTyrPheThrValAlaLeuAsnGluLysPheAspIleHisTyrGluGlnMet 284

RESULT 38
ST1B1_RAT
ID ST1B1_RAT STANDARD; PRT; 299 AA.
AC P52847;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 15-DEC-1998, sequence version 2.
DT 07-FEB-2006, entry version 42.
DE Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-)
DE (Sulfotransferase 1B) (DOPA/tyrosine sulfotransferase).
GN Names:St1b1; Synonyms=St1b1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND PROTEIN SEQUENCE OF 108-117; 126-134
RP AND 260-266.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=96107201; PubMed=8530477; DOI=10.1074/jbc.270.51.30470;
RA Sakakibara Y., Takami Y., Zwieb C., Nakayama T., Suiko M.,

```





OC Cercopithecidae; Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9541;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [MRNA].  
 RC TISSUE=Liver;  
 RA Ogura K., Sateukawa M., Okuda H., Watabe T.;  
 RT "cDNA cloning and bacterial expression of monkey liver phenol  
 sulfotransferase.";  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Catalyzes the sulfate conjugation of catecholamines and  
 phenolic drugs and neurotransmitters (By similarity).  
 CC -1- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + a phenol =  
 adenosine 3',5'-bisphosphate + an aryl sulfate.  
 CC -1- SUBUNIT: Homodimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasm (By similarity).  
 CC -1- SIMILARITY: Belongs to the sulfotransferase family.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; D85514; BAAL2822.1; -; mRNA.  
 DR PIR; G02924; G02924.  
 DR HSP; P50224; 1CJM.  
 DR SMR; P52846; 8-295.  
 DR InterPro; IPR000863; Sulfotransferase.  
 DR Pfam; PF00685; Sulfotransferase; 1.  
 DR ProDom; PD001218; Sulfotransferase; 1.  
 DR Catecholamine metabolism; Lipid metabolism; Steroid metabolism;  
 KW Transferase.  
 KW CHAIN 1 295 Sulfotransferase 1A1.  
 FT FTID=PRO\_0000085129.  
 FT NP\_BIND 48 53 PAPS (By similarity).  
 FT NP\_BIND 130 138 PAPS (By similarity).  
 FT NP\_BIND 193 229 PAPS (By similarity).  
 FT NP\_BIND 257 259 PAPS (By similarity).  
 FT ACT\_SITE 108 108 Proton acceptor (By similarity).  
 SQ SEQUENCE 295 AA; 34239 MW; BD51639D1570A841 CRC64;  
 Alignment Scores:  
 Pred. No.: 1e-22 Length: 295  
 Score: 446.00 Matches: 93  
 Percent Similarity: 54.2% Conservative: 56  
 Best Local Similarity: 33.8% Mismatches: 108  
 Query Match: 10.1% Indels: 18  
 DB: 1 Gaps: 3  
 US-10-768-158-1 (1-2419) x ST1A1\_MACFA (1-295)  
 QY 90 GGCGTGGCGTCCGCGCTTCTGCGGGGAGATGAGGAGATCGCAACTTCCCGTG 149  
 Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheArgAla 36  
 QY 150 CGGCCGACGACGCTGGATCGTCACTTACCCCAAGTCCGACACGAGTTCGCGAG 209  
 Db 37 ArgProAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpValSerGln 56  
 QY 210 GTGCTCTACTTGGTGGAGCGGCGTACCCCGATGAGATCGGCTTGATGAACATCGAC 269  
 Db 57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysArgAlaProIlePhe 76  
 QY 270 GAGCAGTCCCGTCTCGGAGTACCCACAGCGG-----GGCTGGACATCATC 317  
 Db 77 MetArgValProPheLeuGluPheLysValProGlyIleProSerGlyMetGluThrLeu 96  
 QY 318 AAGAACTGACTCTCCCGCTCATCAAGAGCCACCTCGCTTCTTCCCTCT 377  
 Db 97 LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuProGln 116  
 QY 378 GACCTCCCAATGAGACTCCAGGTTCATATATGCTCGCAACCCCAAGGATCTGTG 437  
 Db 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspValAla 136  
 QY 438 GTGCTTATTATCAGTTCCACCGCTCTCTGGGACCATGAGCTACCGAGGACCTTTCAA 497

Db 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp 156  
 QY 498 GAATTTCTGCGGAGGTTTATGAATGATGAAGTGGCTACGGCTCCTGGTTTGACACGTG 557  
 Db 157 SerPheLeuGluLysPheMetAlaGlyGluValSerTyrGlySerTrpTyrGlnHisVal 176  
 QY 558 CAGAGTTCTGGGAGCACCCATGGAGCTGCAAGCTGCTTTTCTCAAGTATGAAGCATG 617  
 Db 177 GlnGluTrpTrpGluLeuSerHisThrHisProValLeuTyrLeuPheTyrGluAspMet 196  
 QY 618 CATCGGACCTGCTGAGATGGTGGAGCAGCTGCCAGATTCTCGGGGGTCTCTGTGAC 677  
 Db 197 LysGluAsnProLysArgGluIleTrpLysIleLeuGluPheValGlyArgSerLeuPro 216  
 QY 678 AAGCCCGAGCTGGAAGCCCTGACGGAGCAGCATGC-----CACCAAGCTGGTGACAC 728  
 Db 217 GluGluThrValAspLeuMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236  
 QY 729 TGTGCAACGCTGAGGCCCTGCCC-----GTG 755  
 Db 237 MetAlaAsnTyrThrThrIleProGlnGluLeuMetAspHisSerIleSerProPheMet 256  
 QY 756 GCGCGGGAAGAGTTGGCTGTGGAAGGACATCTTACCGCTCTCCATGAATGAGAACTTT 815  
 Db 257 ArgLysGlyMetThrGlyAspTrpLysThrPheThrValAlaGlnAsnGluHisPhe 276  
 QY 816 GACTTGGTGTATAACAGAGATGGGAAGTGTGACCTCACGTTT 860  
 Db 277 AspValAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291  
 RESULT 40  
 ST1E1\_BOVIN  
 ID ST1E1\_BOVIN STANDARD; PRT; 295 AA.  
 AC P19217;  
 DT 01-NOV-1990, integrated into UniProtKB/Swiss-Prot.  
 DT 01-NOV-1990, sequence version 55.  
 DE Estrogen sulfotransferase (EC 2.8.2.4) (Sulfotransferase, estrogen-  
 preferring) (ST1E1).  
 GN Name=ST1E1; Synonyms=OST, STE;  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
 OC Pecora; Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [MRNA].  
 RC TISSUE=Placenta;  
 RX MEDLINE=90147277; PubMed=3271383;  
 RA Nash A.R., Glenn W.K., Moore S.S., Kerr J., Thompson A.R.,  
 RA Thompson E.O.P.;  
 RT "Oestrogen sulfotransferase: molecular cloning and sequencing of cDNA  
 for the bovine placental enzyme.";  
 RL Aust. J. Biol. Sci. 41:507-516(1988).  
 RN [2]  
 RP PARTIAL PROTEIN SEQUENCE.  
 RC TISSUE=Placenta;  
 RX MEDLINE=30104087; PubMed=3270501;  
 RA Moore S.S., Thompson E.O.P., Nash A.R.;  
 RT "Oestrogen sulfotransferase: isolation of a high specific activity  
 species from bovine placenta.";  
 RL Aust. J. Biol. Sci. 41:333-341(1988).  
 RN [3]  
 RP PROTEIN SEQUENCE OF 146-160 AND 206-220, AND CHARACTERIZATION.  
 RC TISSUE=Placenta;  
 RX MEDLINE=91152101; PubMed=1900200; DOI=10.1016/0167-4838(91)90279-9;  
 RA Adams J.B.;  
 RT "Enzymic synthesis of steroid sulphates. XVII. On the structure of  
 bovine estrogen sulphotransferase.";  
 RL Biochim. Biophys. Acta 1076:282-288(1991).  
 CC -1- FUNCTION: May control the level of the estrogen receptor by  
 sulfurylating free estradiol.

[illegible]

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzyzanski M.I., Skalska U., Smalhus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [4]  
RP NUCLEOTIDE SEQUENCE OF 1-73.  
RA Falany C.N., Wang J.;  
RT "Mapping of the SULT1B2 gene to human chromosome 4q11-13.";  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: Catalyzes the sulfate conjugation of many hormones,  
CC neurotransmitters, drugs and xenobiotic compounds. Sulfonation  
CC increases the water solubility of most compounds, and therefore  
CC their renal excretion, but it can also result in bioactivation to  
CC form active metabolites. Sulfates dopamine, small phenols such as  
CC 1-naphthol and p-nitrophenol and thyroid hormones, including 3,3'-  
CC diiodothyronine, triiodothyronine, reverse triiodothyronine and  
CC thyroxine.  
CC -1- SUBUNIT: Monomer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasm (By similarity).  
CC -1- TISSUE SPECIFICITY: Highly expressed in the liver, peripheral  
CC blood leukocytes, colon (mucosal lining), small intestine  
CC (jejunum) and spleen. A lesser expression was observed in the  
CC lung, placenta and thymus.  
CC -1- SIMILARITY: Belongs to the sulfotransferase family.  
CC  
CC Copyrighted by the Unifrot Consortium, see <http://www.unifrot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL: D89479; BAA24547.1; -; mRNA.  
DR EMBL: U95726; AAB65154.1; -; mRNA.  
DR EMBL: BC010895; AAH10895.1; -; mRNA.  
DR EMBL: AF184894; AAF05917.1; -; Genomic\_DNA.  
DR PIR: JC5885; JCS885.  
DR PDB: 1XV1; X-ray; A/B=1-296.  
DR Ensembl: ENSG00000173597; Homo sapiens.  
DR HGNC: HGNC:17845; SULT1B1.  
DR MIM: 608436; Gene.  
DR GO: GO:0005829; C:cytosol; NAS.  
DR GO: GO:0008146; F:sulfotransferase activity; IDA.  
DR GO: GO:0018958; P:phenol metabolism; IDA.  
DR GO: GO:0042403; P:thyroid hormone metabolism; IDA.  
DR InterPro: IPR008663; Sulfotransferase.  
DR Pfam: PF00685; Sulfotransfer\_1; 1.  
DR ProDom: PD001218; Sulfotransferase; 1.  
KW 3D-structure; Lipid metabolism; Polymorphism; Steroid metabolism;  
KW Transferase.  
FT CHAIN 1 296 Sulfotransferase family cytosolic 1B  
FT member 1.  
FT /FTId=PRO\_0000085161.  
FT NP\_BIND 48 53 PAPS (By similarity).  
FT NP\_BIND 131 139 PAPS (By similarity).  
FT NP\_BIND 194 230 PAPS (By similarity).  
FT ACT\_SITE 109 109 Proton acceptor (By similarity).  
FT BINDING 259 259 PAPS (By similarity).  
FT VARIANT 145 145 L -> V (in dbSNP:11569736).  
FT /FTId=VAR\_020886.  
FT CONFLICT 183 183 K -> R (in Ref. 3).  
FT CONFLICT 186 186 E -> G (in Ref. 1).  
FT TURN 5 9  
FT STRAND 13 15  
FT TURN 16 17  
FT STRAND 18 20  
FT STRAND 22 24  
FT HELIX 25 26

FT HELIX 28 32  
FT TURN 33 33  
FT TURN 38 39  
FT STRAND 41 46  
FT TURN 47 48  
FT STRAND 49 50  
FT HELIX 51 62  
FT TURN 63 65  
FT TURN 67 68  
FT HELIX 69 71  
FT STRAND 72 73  
FT HELIX 75 78  
FT STRAND 79 79  
FT STRAND 81 81  
FT TURN 82 83  
FT STRAND 86 86  
FT TURN 87 89  
FT STRAND 90 90  
FT HELIX 93 97  
FT TURN 98 99  
FT STRAND 100 100  
FT STRAND 102 103  
FT STRAND 105 108  
FT TURN 112 114  
FT STRAND 115 115  
FT HELIX 117 121  
FT TURN 122 123  
FT STRAND 125 130  
FT HELIX 133 146  
FT STRAND 147 147  
FT TURN 148 149  
FT STRAND 150 150  
FT HELIX 156 164  
FT TURN 165 166  
FT STRAND 168 168  
FT TURN 169 170  
FT HELIX 173 185  
FT TURN 186 186  
FT STRAND 187 187  
FT STRAND 189 193  
FT HELIX 194 199  
FT HELIX 201 211  
FT TURN 212 213  
FT HELIX 218 227  
FT TURN 228 228  
FT STRAND 229 229  
FT HELIX 230 235  
FT TURN 237 239  
FT STRAND 240 241  
FT TURN 242 243  
FT STRAND 244 244  
FT TURN 246 248  
FT STRAND 249 249  
FT TURN 251 253  
FT STRAND 254 254  
FT STRAND 256 257  
FT STRAND 262 263  
FT HELIX 264 267  
FT TURN 268 268  
FT HELIX 271 285  
FT TURN 286 287  
FT STRAND 288 289  
FT STRAND 294 294  
SQ SEQUENCE 296 AA; 34899 MW; AFE61B21BDBD782C CRC64;

Alignment Scores:  
Pred. No.: 1,51e-22  
Score: 443.50  
Percent Similarity: 53.8%  
Best Local Similarity: 33.9%  
Query Match: 10.1%  
DB: 1

Length: 296  
Matches: 94  
Conservative: 55  
Mismatch: 109  
Indels: 19  
Gaps: 3

```

US-10-768-158-1 (1-2419) x ST1B1_HUMAN (1-296)
QY 87 CATGGCGTGGCGTCCCGCCCTTCGCGCGGGAAGATGGAGGAGATCGCAACTTCCTCCG 146
Db ||||| : : : |||||
16 HisGlyTyrProMetThrCysAlaPheAlaSerAsnTrpGluLysIleGluGlnPheHis 35
QY 147 CTGGCGCCAGCGAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 206
Db ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : |||||
36 SerArgProAspAspIleValIleAlaThrTyrProLysSerGlyThrTrpValSer 55
QY 207 GAGTGGTGTACTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 266
Db ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : |||||
56 GluIleIleAspMetIleLeuAsnAspGlyAspIleGluLysCysLysArgGlyPheIle 75
QY 267 GACGAGCAGTCCCGGTCTCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 311
Db ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : |||||
76 ThrGluLysValProMetLeuGluMetThrLeuProGlyLeuArgThrSerGlyIleGlu 95
QY 312 ATCATCAAGGAAGTACCTCTCCCGCTCATCAGAGCCACCTCGCTACCGCTTCTCTG 371
Db ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : |||||
96 GlnLeuGluLysAsnProSerProArgIleValLysThrHisLeuProThrAspLeuLeu 115
QY 372 CCTCTGACCTCCCAATGGAGATCCCAAGGTCTATATGGTTCGCAACCCCAAGGAT 431
Db ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : |||||
116 ProLysSerPheTrpGluAsnAsnCysLysMetIleTyrLeuAlaArgAsnAlaLysAsp 135
QY 432 CTGTGGTGGTCTTATATACGTTCCACCGCTCTCTGCGGACCATGAGCTACCGGACAC 491
Db ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : |||||
136 ValSerValSerTyrTyrHisPheAspLeuMetAsnLeuGlnProPheProGlyThr 155
QY 492 TTTCAGGAAATTCGCGGAGGTTTATGATATGATAGCTGGGCTACGGCTCTGTTGAG 551
Db ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : |||||
156 TrpGluGluTyrLeuGluLysPheLeuThrGlyLysValAlaTyrGlySerTrpPheThr 175
QY 552 CACGTGAGGAGTGTCTGGGACGACCGCATGACGTCGAACGTCCTTTCTCAAGTAA 611
Db ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : |||||
176 HisValLysAsnTrpTrpLysLysLysGluHisProIleLeuPheLeuTyrTrpGlu 195
QY 612 GACATGATCGGACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 671
Db ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : |||||
196 AspMetLysGluAsnProLysGluGluLysLysLysIleIleArgPheLeuGluLysAsn 215
QY 672 TGTGACAGGCCAGCTGGAAGCCTGACGAGCAGCTGCCACACGCTGGTG----- 722
Db ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : |||||
216 LeuAsnAspGluIleLeuAspArgIleIleHisThrSerPheGluValMetLysAsp 235
QY 723 GACCAGTGTGCAAGCTGAGGCTCGCC----- 752
Db ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : |||||
236 AsnProLeuValAsnTyrThrHisLeuProThrThrValMetAspHisSerLysSerPro 255
QY 753 ---GTGGCGCGGGAAGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 809
Db ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : |||||
256 PheMetArgLysGlyThrAlaGlyAspTrpLysAsnTyrPheThrValAlaGlnAsnGlu 275
QY 810 AAGTTTGACTGGTGTATAACACAGATGGGAAGTGTGACCTTCACGTTT 860
Db ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : |||||
276 LysPheAspAlaIleTyrGluThrGluMetSerLysThrAlaLeuGlnPhe 292

RESULT 42
ST1B1_CANFA
ID ST1B1_CANFA STANDARD; PRT; 296 AA.
AC Q95JD5;
DT 04-JAN-2005, integrated into UniProtKB/Swiss-Prot.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 22.
DE Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-)
DE (Sulfotransferase 1B) (CSULT1B1).
GN Name=ST1B1; Synonyms=ST1B2;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
OC Canis.
OX NCBI_taxID=9615;

```

```

[1]
NUCLEOTIDE SEQUENCE [MRNA], FUNCTION, AND TISSUE SPECIFICITY.
RP TISSUE=Liver;
RC MEDLINE=21261348; PubMed=11368519; DOI=10.1006/abbi.2001.2373;
RX Tsou C., Falany C.N., Morgenstern R., Swedmark S.;
RA "Molecular cloning, expression, and characterization of a canine
RT Sulfotransferase that is a human ST1B2 ortholog.";
RL Arch. Biochem. Biophys. 390:87-92(2001).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of many hormones,
CC neurotransmitters, drugs and xenobiotic compounds. Sulfonation
CC increases the water solubility of most compounds, and therefore
CC their renal excretion, but it can also result in bioactivation to
CC form active metabolites. Sulfates thyroid hormones including 3,3'-
CC diiodothyronine, 3,5,3'-triiodothyronine and 3,5'-diiodothyronine
CC (lesser degree) and phenols such as 1-naphthol and p-nitrophenol.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed highly in the colon, kidney and
CC small intestine of male and female dogs. Highly expressed in the
CC jejunum and ileum of the male dog than the female dog, which
CC displayed more expression in duodenum (at protein level).
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL: AY004332; AAF86583.1; -; mRNA.
DR SMR; Q95JD5; 4-296.
DR Ensembl; ENSCAFG0000002871; Canis familiaris.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Lipid metabolism; Steroid metabolism; Transferase.
FT CHAIN 1 296 Sulfotransferase family cytosolic 1B
FT member 1.
FT NP_BIND 48 53 PAPS (By similarity).
FT NP_BIND 131 139 PAPS (By similarity).
FT NP_BIND 194 230 PAPS (By similarity).
FT ACT_SITE 109 109 Proton acceptor (By similarity).
FT BINDING 259 259 PAPS (By similarity).
SQ SEQUENCE 296 AA; 34869 MW; 3D54CF003BF09365 CRC64;

Alignment Scores:
Pred. No.: 1,63e-22 Length: 296
Score: 443.00 Matches: 101
Percent Similarity: 52.4% Conservative: 51
Best Local Similarity: 34.8% Mismatches: 118
Query Match: 10.1% Indels: 20
DB: 1 Gaps: 5

US-10-768-158-1 (1-2419) x ST1B1_CANFA (1-296)
QY 51 ACCCGCGGGAGTTCGAGACCAAGTACTTCGAG---TTCCATGCGGTGCGGCTGCCGCC 107
Db 3 SerProLysAspPheLeuArgLysAsnLeuLysMetIleHisGlyTyrProIleIleTyr 22
QY 108 TTCTGCGCGGGAAGATGGAGGAGATCGCAACTTCCTCCGTCGCGCCAGCAGCTGTGG 167
Db 23 ThrPheAlaAsnAsnTrpGluAsnIleGluGlnPheHisSerArgProAspAspIleIle 42
QY 168 ATCGTCACTACCCCAAGTCCGACACCATCTGCTGCGAGGAGTGCTTACTTGGTGAGC 227
Db 43 IleAlaThrTyrProLysSerGlyThrTrpValSerGluIleValAspMetValLeu 62
QY 228 CAGGCGCTCACCCCGATGATCGGCTTGTATGACATCCAGCAGCAGCTCCCGTCTCTG 287
Db 63 AsnAsnGlyAspValGluLysCysLysArgAspPheIleThrValLysValProMetLeu 82
QY 288 GAGTACCCACAGCGCGGCGCTG-----GACATCATCAAGGAAGTACCTCT 332
Db ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : |||||
83 GluMetAlaValProGlyLeuArgThrSerGlyIleGluGlnLeuGluLysAsnProSer 102

```



```

Db 226 GlnGluThrSerPheGluLysMetLysAspAsnProMetIleAsnArgSerThrValPro 245
Qy 753 -----GTGGCGCGGGGAGAGTTGGGCTGGG 779
Db 246 LysSerIleMetAspGlnSerIleSerProPheMetArgLysGlyThrValGlyAspTrp 265
Qy 780 AAGGACATCTTCACCGTCTCCATGAATGAGAGTTGTGATGGTGATATAACAGAGATG 839
Db 266 LysAsnHisPheThrValAlaGlnSerHisArgLeuAspGluLeuTyrArgLysLysMet 285
Qy 840 GGAAAGTGTGACCTCACGTTT 860
Db 286 GluGlyValSerIleAspPhe 292

RESULT 44
ID Q9XT99_RABIT PRELIMINARY; PRT; 292 AA.
AC Q9XT99;
DT 01-NOV-1999, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 17.
DE Arylsulfotransferase STIA8.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=New Zealand White; TISSUE=Liver;
RX MEDLINE=2109889; PubMed=11181495;
RA Homma W., Kamiyama Y., Yoshinari K., Sasano H., Shimada M., Nagata K.,
RA Yamazoe Y.;
RT "Enzymatic characterization and interspecies difference of phenol
RT sulfotransferases, STIA forms.";
RL Drug Metab. Dispos. 29:274-281(2001).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AB029494; BAA82295.1; -; mRNA.
DR HSSP; P50224; 1CUM.
DR SMR; Q9XT99; 8-289.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 292 AA; 33841 MW; 2BF5F63A259DFFD1 CRC64;

Alignment Scores:
Pred. No.: 1.92e-22 Length: 292
Score: 442.00 Matches: 95
Percent Similarity: 53.9% Conservative: 50
Best Local Similarity: 35.3% Mismatches: 104
Query Match: 10.0% Indels: 20
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x Q9XT99_RABIT (1-292)
Qy 90 GCGCTGCGGCTGCCCGCTTCCTCGCGCGGAGAGTGGAGGATCGCCAACTTCCCGGTG 149
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaMetGlyProLeuGlnSerPheLysAla 36
Qy 150 CGGCCACGACGAGTGTGGATCGTACCTACCCCAAGTCCGCGACGAGTTCTCGCAGGAG 209
Db 37 ArgProAspAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpValSerGlu 56
Qy 210 GTGGTCTACTTGGTGGACCGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
Db 57 IleLeuAspMetIleTyrGlnSerGlyAspGlnGlnLysCysLeuArgAlaProIleTyr 76

```

```

Qy 270 GAGCAGCTCCCGTCTCTGGAGTACCCACAGCCG-----GCCTTGGACATCATC 317
Db 77 IleArgValProPheLeuGluPheLysAlaProSerGlyMetGluThrLeu 96
Qy 318 AAGGAATGACCTTCCCGGCTCATCAAGAGCCACCTGCCCTACCGCTTCTGCCCTCT 377
Db 97 LysAspThrProSerProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116
Qy 378 GACCTCCACAAATGAGACTCCAAGTGCATATATGCTCGCAACCCCAAGATCTGGTG 437
Db 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspValAla 136
Qy 438 GTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGCTACCGAGGACCTTCAA 497
Db 137 ValSerTyrTyrAsnPheTyrHisMetAlaArgValHisProGluProGlyThrTrpAsp 156
Qy 498 GAATTCTCGCGGAGGTTTATGAATGATAAGCTGGCTACGCTCGCTCTGTTTGACACGTG 557
Db 157 SerPheLeuGluLysPheMetAlaGlyLysValSerTyrGlySerTyrTyrGlnHisVal 176
Qy 558 CAGGAGTTCTGGGAGACACCGCATGCACTCGAACCTGCTTTTCTCAAGTATGAAGACATG 617
Db 177 ArgGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheGluAspMet 196
Qy 618 CATCGGACCTGTGTGACGATGGTGGAGCAGCTGGCCAGATTCCTGGGGGTGCTCTGTGAC 677
Db 197 LysGluAsnProLysThrGlnIleLysLysIleLeuGluPheLeuGlyArgSerLeuPro 216
Qy 678 AAGGCCAGCTGGNAGCCCTGACGGAGCAGCTGCCACGAGTGGTGGACCG----- 728
Db 217 GluGluThrValAspArgIleAla---HisCysThrSerPheLysGluMetLysLysAsn 235
Qy 729 ---TGCTGCAACGCTGAGGCCCTGCC----- 752
Db 236 ProMetThrAsnTyrSerThrIleProGluAsnIleMetAspHisAsnValSerProPhe 255
Qy 753 GTGGCGCGGAGAGTGGCTGTGGAAGGACATCTTACCCTGCTCCATGAATGAGAG 812
Db 256 MetArgArgGlyValAlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArg 275
Qy 813 TTTGACTTGGTGATATAAACAGAGATG 839
Db 276 PheGluAlaAspTyrAlaGluLysMet 284

RESULT 45
Q6GP49_XENLA PRELIMINARY; PRT; 297 AA.
ID Q6GP49_XENLA PRELIMINARY; PRT; 297 AA.
AC Q6GP49;
DT 19-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 14.
DE MGC80677 protein.
GN Name=MGC80677;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uudin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

```





DB:	2	Gaps:	4	
US-10-768-158-1 (1-2419) x Q3T0Y3_BOVIN (1-296)				
QY	87	CATGGCGTGGCGCTCCCGCCCTTCGCGCGGGAAGATGGAGGAGATCGCAACTTCCCG	146	
Db	16	HisGlyCysProIleThrTyraIaPheAlaAsnAsnTrpGluLysIleGluGlnPheGln	35	
QY	147	GTGGCGCCAGCGACGTGTGGATCGTACCTACCCAAAGTCGGGACAGCTTGTGCGAG	206	
Db	36	SerArgProAspAspIleMetIleValThrTyProLysSerGlyThrThrTrrpIleSer	55	
QY	207	GAGTGTGTACTTGGTGAGCGGCGCTGACCCCGATGATCGGCTTATGACATC	266	
Db	56	GluIleValAspMetMetValLeuHisAspGlyAspValGluLysCysLysArgAspValIle	75	
QY	267	GACGAGCAGCTCCCGCTCTGAGTACCCACAGCCG-----GGCTGGAC	311	
Db	76	ThrAlaLysValProMetLeuGluLeuAlaLeuProGlyLeuArgThrSerGlyLeuGlu	95	
QY	312	ATCATCAAGGAATGACCTCTCCCGCTCATCAAGAGCCACTGCTCCCTACCGCTTTCTG	371	
Db	96	GlnLeuGluLysAsnProSerProArgValLysThrHisLeuProIleAspLeuIle	115	
QY	372	CCCTCTGACCTCCACAAATGGAGACTCCAAAGTCATCTATATGGTCGCAACCCAGGAT	431	
Db	116	ProLysSerPheTrpGluAsnAsnCysLysIleIleTyLeuAlaArgAsnAlaLysAsp	135	
QY	432	CTGTGGTGCTTATTATCAGTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGGCACC	491	
Db	136	ValAlaValSerPheTyHisPheAspLeuMetAsnLeuGlnProLeuProGlyThr	155	
QY	492	TTTCAAGAAATCTGCGGAGGTTTATGAATGATAAGCTGGGCTCGGCTCTCTGGTTTCAG	551	
Db	156	TrpGlyGluTyLeuGluLysPheLeuThrGlyAsnValAlaTyGlySerTrpPheAsn	175	
QY	552	CACGTGACGAGTCTGGGACACCGCATGACCTGCAACGCTGCTTTCTCAAGTATGAA	611	
Db	176	HisValLysSerTrpTrpLysLysGluGlyHisProIleLeuPheLeuPheTyGlu	195	
QY	612	GACATGATCGGACCTGTCAGCATGTGTGAGCAGCTGGCAGATTCCTGGGGTGTC	671	
Db	196	AspMetLysGluAsnProLysGlnGluIleLysLysValValArgPheLeuGluLysAsn	215	
QY	672	TGTGACAAAGGCCAGCTGGAAGCCCTGACGAGCACTGC-----CACCAGCTGTGGAC	725	
Db	216	LeuAspAspGluIleLeuAspLysIleIleTyHisThrSerPheGluMetMetLysAsp	235	
QY	726	CAG---TGCTGCAAGCTGAGCCCTGCCC-----	752	
Db	236	AsnProLeuValAsnTyThrHisLeuProSerGluValMetAspHisSerLysSerSer	255	
QY	753	---GTGGCGCGGGAAGAGTGGGCTGGGAGGACATCTTCACGCTCCCATGATGAG	809	
Db	256	PheMetArgLysGlyIleAlaGlyAspTrpLysAsnTyPheThrValAlaGlnAsnGlu	275	
QY	810	AAGTTTGACTTGGTGATATAACAGAAAGTGTGGAAGTGTGCTCACCTCACCTTT	860	
Db	276	LysPheAspAlaIleTyLysLysGluMetSerGluThrGluLeuGlnPhe	292	
RESULT 47				
STI2 RAT				
ID	STI2 RAT	STANDARD;	PRT; 295 AA.	
AC	P52845;			
DT	01-OCT-1996,	integrated into UniProtKB/Swiss-Prot.		
DT	01-OCT-1996,	sequence version 1.		
DE	07-FEB-2006,	entry version 38.		
DE	Estrogen sulfotransferase, isoform 2 (EC 2.8.2.4) (EST-2)			
DE	(Sulfotransferase, estrogen-prefering) (Estrogen sulfotransferase).			
GN	Name=Ste2; Synonyms=Ste;			
OS	Rattus norvegicus (Rat);			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;			

OC	Muroidea; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE [MRNA].		
RC	STRAIN=Fischer 344; TISSUE=Liver;		
RX	MEDLINE=96305357; PubMed=8688469; DOI=10.1016/0167-4781(96)00065-6;		
RA	Rikke B.A.; Roy A.K.;		
RT	"Structural relationships among members of the mammalian		
RL	sulfotransferase gene family.";		
RL	Biochim. Biophys. Acta 1307:331-338(1996).		
CC	-!- FUNCTION: Sulfation of estrone and estradiol. May control the		
CC	level of the estrogen receptor by sulfurylating free estradiol (By		
CC	similarity).		
CC	-!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =		
CC	adenosine 3',5'-bisphosphate + estrone 3-sulfate.		
CC	-!- SUBUNIT: Homodimer (By similarity).		
CC	-!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).		
CC	-!- SIMILARITY: Belongs to the sulfotransferase family.		
CC	-----		
CC	Copyrighted by the UniProt Consortium, see <a href="http://www.uniprot.org/terms">http://www.uniprot.org/terms</a>		
CC	Distributed under the Creative Commons Attribution-NoDerivs License		
CC	-----		
DR	EMBL; U50205; AAB07681.1; -; mRNA.		
DR	HSSP; P49891; 1B06.		
DR	SMR; P52845; 7-294.		
DR	Ensembl; ENSRNOG00000001957; Rattus norvegicus.		
DR	RGD; 3776; Ste.		
DR	InterPro; IPR000863; Sulfotransferase.		
DR	Pfam; PF00685; Sulfotransferase 1; 1.		
DR	ProDom; PD001218; Sulfotransferase; 1.		
KW	Lipid-binding; Steroid-binding; Transferase.		
FT	CHAIN 1 295 Estrogen sulfotransferase, isoform 2.		
FT	NP_BIND 48 53 PAPS (By similarity).		
FT	NP_BIND 130 138 PAPS (By similarity).		
FT	NP_BIND 193 229 PAPS (By similarity).		
FT	NP_BIND 257 259 PAPS (By similarity).		
FT	ACT_SITE 108 108 Proton acceptor (By similarity).		
SQ	SEQUENCE 295 AA; 35365 MW; 149B5C9D46039AAF CRC64;		
Alignment Scores:			
Pred. No.:	6.03e-22	Length:	295
Score:	435.00	Matches:	92
Percent Similarity:	53.3%	Conservative:	60
Best Local Similarity:	32.3%	Mismatches:	115
Query Match:	9.9%	Indels:	18
DB:	1	Gaps:	3
US-10-768-158-1 (1-2419) x STI2_RAT (1-295)			
QY	60	GAGTTCGAGAGCAAGTACTTCGAGTTCATGCGGTGCGGCTGCGCCCTTCTGCGCGGG	119
DB	7	GluTyTrpGluValPheGlyAspPheHisGlyPheLeuMetAspLysLeuPheThrLys	26
QY	120	AAGATGAGGAGATCGCCAACTTCCGGTGGCGCCAGCGACGTGTGGATCGTCACCTAC	179
DB	27	TyrTrpGluAspValGluThrPheSerAlaArgProAspAspLeuLeuValThrTyr	46
QY	180	CCCAAGTCCGGACCACTGCTGTCAGAGGTGCTTACTTGGTGAGCCAGGCGCGTGAC	239
DB	47	ProLysSerGlySerThrTrpIleGlyGluIleValAspMetIleTyLysGluGlyAsp	66
QY	240	CCCGATGAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGGTCTCGGAGTACCCACAG	299
DB	67	ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyLeuGluCysArgAsn	86
QY	300	CCG-----GGCTGGAGCATCATCAAGAACTGACCTTCCCGCGCTCATCAAG	347
DB	87	GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleValLys	106
QY	348	AGCCACTGCGCTACCGCTTCTGCGCTCTGACCTCCACCAATGGAGACTCCAAGGTATC	407
DB	107	ThrHisLeuProAlaLysLeuLeuProAlaSerPheTrpGluLysAsnCysLysIleIle	126

QY 408 TATATGGCTCGCACCCACAGGATCTGGTGTCTTATATATCATGTTCCACCGCTCTCTG 467  
 Db 127 TyrLeuCysArgAnaAlaLysAspValValSerTyrTyrPhePheLeuLeuMet 146  
 QY 468 CGGACCATGAGTACCGAGGACCTTCAAGAAATCTCGCGAGGTTTATGAATGATAAG 527  
 Db 147 LysSerTyrProAnProLysSerPheSerGluPheValGluLysPheMetGluGluGln 166  
 QY 528 CTGGGCTACGGCTCTCTGGTTGACACGTGCGAGGTTCTGGGAGCCCGCATGACTCG 587  
 Db 167 ValProTyrGlySerTyrAspHisValLysSerTrpTrpGluLysSerLysAsnSer 186  
 QY 588 AACGTGCTTTTCTCAAGTATGAACATGATCGGACCTGTCGATGATGGTGACGAC 647  
 Db 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspTyrArgGluValValLys 206  
 QY 648 CTGGCCAGATTCTCGGGGTCTCTGTGACAAGCCCGAGCTGGAAGCCCTGACGAGCAC 707  
 Db 207 LeuLeuGluPheLeuGluArgAspProSerAlaGluLeuValAspArgGlieGlnHis 226  
 QY 708 TGC-----CACACAGCTGTGGACAGTGTGCAACGCTGAGGCCCTGCC----- 752  
 Db 227 ThrSerPheGlnGluMetLysAsnAsnProCysThrAsnTyrSerMetLeuProGluThr 246  
 QY 753 -----GTGGCCCGGGAGAGTTGGCTGTGGAAGGAC 785  
 Db 247 MetLeuAspLeuLysValSerProPheMetArgLysGlyLeuValGlyAspTrpLysAsn 266  
 QY 786 ATCTTCACCGCTCTCCATGAATGAGAAGTTTGACTGGTGTATATAACAGAGATGGGAAG 845  
 Db 267 HisPheProGluAlaLeuArgGluArgPheGluGluHisTyrGlnGlnMetLysAsp 286  
 QY 846 TGTGACCTCAGTTT 860  
 Db 287 CysProValLysPhe 291

RESULT 48  
 ID STIE6\_RAT  
 AC P49890; STANDARD; PRT; 295 AA.  
 DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.  
 DT 01-OCT-1996, sequence version 1.  
 DT 07-FEB-2006, entry version 37.  
 DE Estrogen sulfotransferase, isoform 6 (EC 2.8.2.4) (EST-6)  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [MRNA].  
 RC TISSUE=Liver;  
 RX MEDLINE=95161323; PubMed=7857871; DOI=10.1016/0960-0760(94)00147-5;  
 RA Falany J.L., Krasnykh V., Mikheeva G., Falany C.N.;  
 RT "Isolation and expression of an isoform of rat estrogen  
 sulfotransferase.";  
 RL J. Steroid Biochem. Mol. Biol. 52:35-44 (1995).  
 CC -1- FUNCTION: Sulfation of estrone and estradiol. May control the  
 CC level of the estrogen receptor by sulfurylating free estradiol.  
 CC -1- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =  
 CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.  
 CC -1- SUBUNIT: Homodimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasm.  
 CC -1- DEVELOPMENTAL STAGE: Expressed in males rats, but not in females.  
 CC -1- INDUCTION: Induced by androgens and suppressed by estrogens.  
 CC Expression is under the influence of pituitary growth hormone and  
 CC thyroid hormone.  
 CC -1- SIMILARITY: Belongs to the sulfotransferase family.  
 CC  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----  
 DR EMBL; S76490; AAB33442.1; -; mRNA.  
 DR HSP; P49891; IBO6.  
 DR SMR; P49890; 7-294.  
 DR Ensembl; ENSRNOG00000001957; Rattus norvegicus.  
 DR InterPro; IPR000863; Sulfotransferase.  
 DR Pfam; PF00685; Sulfotransferase; 1.  
 DR ProDom; PD001218; Sulfotransferase; 1.  
 KW Lipid-binding; Steroid-binding; Transferase.  
 FT CHAIN 1 295 Estrogen sulfotransferase, isoform 6.  
 FT FTID=PRO 000085158.  
 FT NP\_BIND 48 53 PAPS (By similarity).  
 FT NP\_BIND 130 138 PAPS (By similarity).  
 FT NP\_BIND 193 229 PAPS (By similarity).  
 FT NP\_BIND 257 259 PAPS (By similarity).  
 FT ACT\_SITE 108 108 Proton acceptor (By similarity).  
 SQ SEQUENCE 295 AA; 35303 MW; 9FBD5861AFDC9B05 CRC64;  
 Alignment Scores:  
 Pred. No.: 6,03e-22 Length: 295  
 Score: 435.00 Matches: 93  
 Percent Similarity: 53.1% Conservative: 62  
 Best Local Similarity: 31.8% Mismatches: 105  
 Query Match: 9.9% Indels: 32  
 DB: 1 Gaps: 4  
 US-10-768-158-1 (1-2419) x STIE6\_RAT (1-295)  
 QY 60 GAGTTCGAGACGAAGTACTTCGAGTTCCATGGCGTGGCTGCCGCCCTTCTCGCGGG 119  
 Db 7 GluTyrTyrGluValPheGlyAspPheHisGlyValLeuValAspLysLeuPheThrLys 26  
 QY 120 AAGATGAGGAGATCGCCAACTCCCGGTGCGGCCCGACGACGTGGATCGTCACCTAC 179  
 Db 27 TyrTrpGluAspValGluThrPheSerAlaArgProAspAspLeuValThrTyr 46  
 QY 180 CCCAAGTCCGGCACCGACTTGTGTCAGGAGGTGTCTACTTGTGAGCGGCGGTGAC 239  
 Db 47 ProLysSerGlySerThrTrpIleGlyLeuValAspMetIleTyrLysGluGlyAsp 66  
 QY 240 CCGATGAGATCGGCTTGATGAACATCGACGACGATCCCGTCTCTGGAGTACCCACAG 299  
 Db 67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86  
 QY 300 CCG-----GGCTGGACATCATCAAGAACTGACCTCCCGCCTCATCAAG 347  
 Db 87 GluAspLeuLeuAsnGlyIleLysGlnLeuLysGluSerProArgIleValLys 106  
 QY 348 AGCACCTGCGCTACCGCTTCTGCGCTCTGACCTCCACAATGGAGACTCCAAAGGTGATC 407  
 Db 107 ThrHisLeuProAlaLysLeuLeuProAlaSerPheTrpGluLysAsnCysLysIleIle 126  
 QY 408 TATATGGCTCGCAACCCCAAGGATCTGGTGTCTTATTATCATGTTCCACCGCTCTCTG 467  
 Db 127 TyrLeuCysArgAsnAlaLysAspValValSerTyrTyrPhePheLeuIleIle 146  
 QY 468 CGGACCATGAGTACCGAGGACCTTCAAGAAATCTCGCGAGGTTTATGAATGATAAG 527  
 Db 147 LysSerTyrProAnProLysSerPheSerGluPheValGluLysPheMetGluGluGln 166  
 QY 528 CTGGGCTACGGCTCTCTGGTTGACGACGTGCGAGGTTCTGGGAGCCCGCATGACTCG 587  
 Db 167 ValProTyrGlySerTyrAspHisValLysSerTrpTrpGluLysSerLysAsnSer 186  
 QY 588 AACGTGCTTTTCTCAAGTATGAACATGATCGGACCTGTCGATGATGGTGACGAC 647  
 Db 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspTyrArgGluValValLys 206  
 QY 648 CTGGCCAGATTCTCGGGGTCTCTGTGACAAGCCCGAGCTGGAAGCCCTGACGAGCAC 707  
 Db 207 LeuLeuGluPheLeu 226  
 QY 708 TGCCACCGACGCTGGTG-----GACCATGCTGCAAC 737

Db 220 ValAspLysIleIleGlnHisThrSerPheGlnGluMetLysAsnAsnProCysThrAsn 239  
FT ::::::::::: :::: ||||  
FT NP\_BIND 51 56 /FTid=PRO\_0000085173.  
FT PAPS (By similarity).  
FT NP\_BIND 135 143 PAPS (By similarity).  
FT NP\_BIND 199 235 PAPS (By similarity).  
FT NP\_BIND 261 263 PAPS (By similarity).  
FT ACT\_SITE 113 113 Proton acceptor (By similarity).  
FT CONFLICT 193 193 I -> M (in Ref. 1).  
FT CONFLICT 193 193 N -> T (in Ref. 1).  
SQ SEQUENCE 299 AA; 35035 MW; 805D82915B45EB6A CRC64;  
  
Alignment Scores:  
Pred. No.: 6.57e-22 Length: 299  
Score: 434.50 Matches: 101  
Percent Similarity: 54.5% Conservative: 51  
Best Local Similarity: 36.1% Mismatches: 109  
Query Match: 9.9% Indels: 19  
DB: 1 Gaps: 6  
  
US-10-768-158-1 (1-2419) x ST1S1\_BRARE (1-299)  
  
Qy 78 TTCGAGTTCATCGCGTGGCGCTCCGCGCTTCGCGCGGGAAGATGGAGAGATCGCC 137  
||| ||||| ||||| ::::  
Db 16 PheGluPheGluGlyIleSerMetIleAsnHisPheThrGluAsnTrpGluLysVallys 35  
  
Qy 138 AACTTCCCGTGGCGCCGACGAGCGTGTGGATCGTACCTACCCCAAGTCCGGCCACG 197  
||| ||||| ||||| ::::  
Db 36 AsnPheGlnAlaArgProAspAspIleLeuIleAlaThrTyProLysAlaGlyThrThr 55  
  
Qy 198 TTGCTGCAGGAGGTGCTACTGCTGAGCGAGCGCGCTGAC--CCCGATGAGATCGCG 254  
||| ||||| ||||| ::::  
Db 56 TrpValSerTyIleLeuAspLeuLeuTyPheGlyGluAsnAlaProGluGluHisThr 75  
  
Qy 255 TTGATGAACATCGACGAGCTCCGCGTCTCGGAG-----TACCCACAGCGG 302  
||| ||||| ||||| ::::  
Db 76 SerGlnProIleTyMetArgValProPheLeuGluSerCysPheLysValIleAlaSer 95  
  
Qy 303 GGCTTGACATCATCAAGAACTG---ACCTCTCCCGCCTCATCAAGACGACCTGCC 359  
||| ||||| ||||| ::::  
Db 96 GlyThrGluLeuAlaAspAsnMetThrThrSerProArgLeuIleLysHisLeuPro 115  
  
Qy 360 TACCGCTTTCGCGCTCTGACCTCCACAATGGAGACTCCAAGTCCATCTATATGCTCGC 419  
||| ||||| ||||| ::::  
Db 116 ValGlnLeuIleProLysSerPheTrpGluGlnAsnSerArgValTyValAlaArg 135  
  
Qy 420 AACCCCAAGGATCGTGGTGTCTTATATCAGTTTCCACCGCTCTCTCGGACCATGAGC 479  
||| ||||| ||||| ::::  
Db 136 AsnAlaLysAspAsnValIleSerTyPheHisPheAspArgMetAsnIleValGluPro 155  
  
Qy 480 TACGAGGCACCTTCAAGAAATTCGCGGAGGTTTATGAATGATTAAGCTGGGCTACGCG 539  
||| ||||| ||||| ::::  
Db 156 AspProGlyAspTrpAsnThrPheLeuHisArgPheMetAspGlyLysSerValPheGly 175  
  
Qy 540 TCCTGGTTTGAGCAGCTGACGAGTCTCTGGGAGCACGCGCATGGAC---TCGAACGTGTT 596  
||| ||||| ||||| ::::  
Db 176 ProTrpTyArgHisValAsnGlyTyTrpGluLysLysGlnThrTySerAsnLeuLeu 195  
  
Qy 597 TTTCTCAAGTATGAAGACATCATCGGACCTCGGTGTGACGATGGTGGAGCAGCTGCCAGA 656  
||| ||||| ||||| ::::  
Db 196 TyrLeuPheTyGluAspLeuValGluAspThrGlyArgGluValAspArgLeuCysSer 215  
  
Qy 657 TTCTTGGGGGTCTCTGTGACAGCCGACGCTGGAAGCCCTGAGGAGACATCTGCCAC--- 713  
||| ||||| ||||| ::::  
Db 216 PheLeuGlyLeuSerThrSerValSerAspArgGluLysIleThrLysAspValGlnPhe 235  
  
Qy 714 -----CAGCTGGTGGACCATGCTGCAACGCTCAGGCGCCCTGCCCGTG----- 755  
||| ||||| ||||| ::::  
Db 236 AspAlaMetLysGlnAsnLysMetThrAsnTySerThrLeuProValMetAspPheLys 255  
  
Qy 756 -----GGCCGGGAAGAGTTGGGCTGTGGAGAGCATCTTCCACGCTCTCC 800  
||| ||||| ||||| ::::  
Db 256 IleSerProPheMetArgLysGlyValGlyAspTrpLysAsnHisPheThrValAla 275  
  
Qy 801 ATGAATGAGAAGTTTGACTTGGTGTATAAACAGAAAGATGGGAAAGTGTGACCTCAGCTTT 860  
||| ||||| ||||| ::::

Db 220 ValAspLysIleIleGlnHisThrSerPheGlnGluMetLysAsnAsnProCysThrAsn 239  
Qy 738 GCTGAGGCCCTGCC-----GTGGCGCGGGA 764  
Db 240 TyrSerMetLeuProGluThrMetIleAspLeuLysValSerProPheMetArgLysGly 259  
Qy 765 AGAGTTGGCTGTGGAAGACATCTTACCGCTCTCCATGAATGAGAAAGTTTGAAGTTGGTG 824  
Db 260 IleValGlyAspTrpArgAsnHisPheProGluAlaLeuArgGluArgPheGluGluHis 279  
Qy 825 TATAACAGAGATGGGAAGTGTGACCTCAGCTTT 860  
Db 280 TyrGlnArgHisMetLysAspCysProValThrPhe 291  
  
RESULT 49  
ST1S1\_BRARE  
ID ST1S1\_BRARE STANDARD; PRT; 299 AA.  
AC O6PH37; Q7T1C8;  
DT 31-AUG-2004, integrated into UniProtKB/Swiss-Prot.  
DT 05-JUL-2004, sequence version 1.  
DT 07-FEB-2006, entry version 15.  
DE Cytosolic sulfotransferase 1 (EC 2.8.2.-) (SULT1 ST1).  
GN Name=sult1st1;  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [MRNA], ACTIVITY, TISSUE SPECIFICITY, AND ENZYME REGULATION.  
RX MEDLINE=22640919; PubMed=12755695;  
RX DOI=10.1046/j.1432-1033.2003.03608.x;  
RA Sugahara T., Liu C.-C., Pai T.G., Colloidi P., Suiko M., Sakakibara Y., Nishiyama K., Liu M.-C.,  
RT "Sulfation of hydroxychlorobiphenyls. Molecular cloning, expression, and functional characterization of zebrafish SULT1 sulfotransferases.";  
RL Eur. J. Biochem. 270:2404-2411(2003).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC TISSUE=kidney;  
RG NIH - Zebrafish Gene Collection (ZGC) project;  
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Catalyzes the conjugation of sulfate to a variety of xenobiotic and endogenous compounds, including 2-naphthol, hydroxychlorobiphenyls, dopamine and T3 (triiodo-L-thyronine).  
CC -!- ENZYME REGULATION: Inhibited by Co(2+), Zn(2+), Cd(2+) and Pb(2+) ions. Inactivated by Hg(2+) and Cu(2+) ions.  
CC -!- BIOPHYSICOCHEMICAL PROPERTIES:  
CC pH dependence:  
CC Optimum pH is 6.0-9.0;  
CC Temperature dependence:  
CC Thermostable from 20 to 43 degrees Celsius;  
CC -!- SUBCELLULAR LOCATION: Cytoplasm.  
CC -!- TISSUE SPECIFICITY: Expressed in liver.  
CC -!- SIMILARITY: Belongs to the sulfotransferase family.  
CC  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NonDerivs License  
CC  
CC EMBL; AY181064; AAO64983.1; -; mRNA.  
DR EMBL; BC056729; AAH56729.1; -; mRNA.  
DR ZFIN; ZDB-GENE-030131-2144; sult1st1.  
DR GO; GO:0005737; C:cytoplasm; NAS.  
DR GO; GO:0008146; F:sulfotransferase activity; IDA.  
DR GO; GO:0006805; P:xenobiotic metabolism; IDA.  
DR InterPro; IPR000863; Sulfotransferase.  
DR Pfam; PF00685; Sulfotransferase 1.  
DR ProDom; PD001218; Sulfotransferase; 1.  
KW Catecholamine metabolism; transferase.  
FT CHAIN 1 299 Cytosolic sulfotransferase 1.



Qy	120	AAGATGGAGGAGATCGCAACTTCCCGTGGCGGCCAGCGACGTGTGGATCTGCACCTAC	179
		::: :::	
Db	26	TyrTrpAspAsnValGluAlaPheGlnAlaArgProAspAspLeuValIleAlaThrTyr	45
Qy	180	CCCAAGTCGGCACCAGCTTCTGCAGCAGAGGTGTCTACTTGGTGAGCCAGCGCGCTGCAC	239
		:::	
Db	46	ProLysSerGlyThrThrTrpValSerGluIleValTyrMetIleTyrLysGluGlyAsp	65
Qy	240	CCCAGATGAGATCGCGTTTGATCAACATCCAGCAGCAGCTCCCGGTCTCTGGAGTACCCACAG	299
		::: :::	
Db	66	ValGluLysCysLysGluAspValIlePheAsnArgIleProPheLeuGluCysArgLys	85
Qy	300	CCG-----GGCTGGACATCATCAAGGAACATGACCTCTCCCGCCTCATCAAG	347
		: :::	
Db	86	GluAsnLeuMetAsnGlyValLysGlnLeuAspGluMetAsnSerProArgIleValLys	105
Qy	348	AGCCACCTGCCTACCGCTTCTGCCCTCTCACCTCCACATGGAGACTCCCAAGGTCATC	407
		:::	
Db	106	ThrHisLeuProGluLeuLeuLeuProAlaSerPheTrpGluLysAspCysLysIleIle	125
Qy	408	TATATGGTCGCAACCCCAAGCATCTGTGTGTCTTATTATCATGTTCCACCGCTCTCTG	467
		: :::	
Db	126	TyrLeuCysArgAsnAlaLysAspValAlaValSerPheTyrTyrPhePheLeuMetVal	145
Qy	468	CGGACCATGAGTACCGAGGACACCTTTCAAGAATTCTCCGGAGTCTTATGAATGATAAG	527
		: :::	
Db	146	AlaGlyHisProAsnProGlySerPheProGluPheValGluLysPheMetGlnGlyGln	165
Qy	528	CTGGGCTACGGCTCTGTTTGACGACGTGCAGGAGTCTCTGGGAGCACCAGCATGGACTCG	587
		:::       :::	
Db	166	ValProTyrGlySerTrpTyrLysHisValLysSerTrpTrpGluLysGlyLysSerPro	185
Qy	588	AACGTGCTTTTCTCAAGTAGAAGACATGATCGGGACCTGTGTACGATGTGGAGCAG	647
		:::	
Db	186	ArgValLeuPheLeuPheTyrGluAspLeuLysGluAspIleArgLysGluValIleLys	205
Qy	648	CTGCCACAGATCTCTGGGGGTCTCTGTGACAAAGCCCGACGTGGAAAGCCCTGACGGAGCAC	707
		:::	
Db	206	LeuIleHisPheLeuGluArgLysProSerGluGluLeuValAspArgIleIleHisHis	225
Qy	708	TGCCAC-----CAGCTGTGGACCAAG	728
		:::	
Db	226	ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrThrLeuProAspGlu	245
Qy	729	TGCTGCAACGCTGAGGCCCTGCC--CTGGCGCGGGAAGAGTTGGGTGTGGGAAGGAC	785
		:::	
Db	246	IleMetAsnGlnLysLeuSerProPheMetArgLysGlyIleThrGlyAspTrpLysAsn	265
Qy	786	ATCTTCCACCGTCTCCATGAATGAGAAGTTTGACTTGGTGTATAAACAGAAGATGGGAAG	845
		:::	
Db	266	HisPheThrValAlaLeuAsnGluLysPheAspLysHisTyrGluGlnGlnMetLysGlu	285
Qy	846	TGTGACCTCAGTTT	860
Db	286	SerThrLeuLysPhe	290
RESULT 51			
Q53X91_HUMAN			
ID	Q53X91_HUMAN	PRELIMINARY;	PRT; 294 AA.
AC	Q53X91;		
DT	24-MAY-2005,	integrated into UniProtKB/TrEMBL.	
DT	24-MAY-2005,	sequence version 1.	
DT	07-FEB-2006,	entry version 3.	
DE	SULT1B1 protein (Fragment).		
GN	Name=SULT1B1;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa;		
OC	Mammalia; Eutheria;		
OC	Primates; Catarrhini;		
OC	Hominoidea;		
OX	NCBI_TaxID=9606;		
RP	{1}		
RP	NUCLEOTIDE SEQUENCE.		

Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;  
 "Cloning of human full open reading frames in Gateway (TM) system entry  
 vector (pDONR201).";  
 RT Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
 RL  
 CC  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC  
 DR EMBL; CR407621; CAG28549.1; -; mRNA.  
 DR SMR; Q53X91; 3-292.  
 DR Ensembl; ENSG0000109193; Homo sapiens.  
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.  
 DR InterPro; IPR000863; Sulfotransferase.  
 DR Pfam; PF06885; Sulfotransferase\_1; 1.  
 DR ProDom; PD001218; Sulfotransferase; 1.  
 DR NON\_TER 294 294  
 SQ SEQUENCE 294 AA; 35126 MW; 9EC8923D20757D57 CRC64;

Alignment Scores:

Pred. No.:	7,1e-22	Length:	294
Score:	424.00	Matches:	96
Percent Similarity:	53.0%	Conservative:	55
Best Local Similarity:	33.3%	Mismatches:	116
Query Match:	9.9%	Indels:	18
DB:	2	Gaps:	3

US-10-768-158-1 (1-2419) x Q53X91\_HUMAN (1-294)

Qy	60	GAGTTCGAGGACGAGTACTTCGAGTTCATCGGTGGCGGTGGCGCCCTTCTGCGCGGG	119
Db	6	AspTyrTyrGluLysPheGluValHisGlyIleLeuMetTyrLysAspPheValLys	25
Qy	120	AAGATCGAGGAGATGCCAACTTCGCGTGGCGCCAGCGAGTGTGATCGTCACTAC	179
Db	26	TyrTrpAsnValGluAlaPheGlnAlaArgProAspPheValLysAlaThrTyr	45
Qy	180	CCCAAGTCGCGCACCATCTTCGAGGAGTGTCTACTTTGGTGGACGAGGCGCTGAC	239
Db	46	ProLysSerGlyThrThrTrpValSerGluIleValTyrMetIleTyrLysGluGlyAsp	65
Qy	240	CCCGATCGAGATCGGCTTGATGAACATCGACGAGCATCCCGTCTGGAGTACCCACAG	299
Db	66	ValGluLysCysLysGluAspValIlePheAsnArgIleProPheLeuGluCysArgLys	85
Qy	300	CCG-----GGCCTGGACATCATCAAGAACTGACCTCTCCCGCCCTCATCAAG	347
Db	86	GluAsnLeuMetAsnGlyValLysGlnLeuAspGluMetAsnSerProArgIleValLys	105
Qy	348	AGCACCTCGCTACCGCTTCTCGCTCTGACCTCCACATGAGACTCAAGTTCATC	407
Db	106	ThrHisLeuProGluLeuLeuProAlaSerPheTrpGluLysAspCysLysIleIle	125
Qy	408	TATATGGCTCGCAACCCCAAGGATCTGGTGGTGTCTTATTATCAGTTCCACCGCTCTCTG	467
Db	126	TyrLeuCysArgAsnAlaLysAspValAlaValSerPheTyrTyrPhePheLeuMetVal	145
Qy	468	CGGACCATGAGCTACCGAGGCACCTTTCAAGAATTTGCGCGAGGTTTATGAATGATAG	527
Db	146	AlaGlyHisProAsnProGlySerPheProGluPheValGluLysPheMetGlnGln	165
Qy	528	CTGGGCTACGGTCTCGGTTTTCAGACATGAGGAGTTCGCGAGCATCCCGCTCTCTG	587
Db	166	ValProTyrGlySerTrpTyrHisValLysSerTrpTrpGluLysGlyLysSerPro	185
Qy	588	AACGTCCTTTTCTCAAGTATGAAGCATGTCATCGGACCTCGGTGACGATGGTGGAGCAG	647
Db	186	ArgValLeuPheLeuPheTyrGluAspLeuLysGluAspIleArgLysGluValIleLys	205
Qy	648	CTGCCCAGATTCTGGGGGTGCTCTGTCACAGGCCAGCTGGGAAGCCCTGACGGAGCAG	707
Db	206	LeuIleHisPheLeuGluArgLysProSerGluGluLeuValAspArgIleIleHisHis	225
Qy	708	TGCCAC-----CAGCTGGTGGACGAG	728









```

Db      198 LysValMetArgPheLeuAspLysAspLeuSerGluGluValLeuGluLysValHis 217
      ::::  |||||  :::  |||||
Qy      696 CTGACGGAGCACTGCCACCACCTGGTGACAGCTGCTCAACGCTGAGGCCCTGCC--- 752
      ||:::  ||||  :::  ||||
Db      218 LeuSerPheAspHisMetLysAspAsnProMetAlaAsnPheSerAlaPheProSer 237
      ::::  ||||  :::  ||||
Qy      753 -----GTGGCGCGGGAAGAGTTGGGCTGTGGGAAG 782
      ::::  ||||  :::  ||||
Db      238 AspValValAspGlnSerGlnTyrLysPheMetArgLysGlyLysValGlyAspTrpLys 257
      ::::  ||||  :::  ||||
Qy      783 GACATCTTCACCGTCTCCATCAATGAGAGTTGCTGGTGTATTAACAGAGAAGTGGGA 842
      |||||  |||||  |||||  |||||
Db      258 SerHisPheThrValGlnGlnAsnGluMetPheGluGluLysTyrGlnGlnMetHis 277
      ::::  ||||  :::  ||||
Qy      843 AAGTGTGACCTCAGCTTGTGACTTT 866
      :::  |||  :::
Db      278 GlySerAlaMetLysPheArgTyr 285
      :::  |||  :::

RESULT 55
STIE1_MOUSE
ID STIE1_MOUSE STANDARD; PRT; 295 AA.
AC P49891;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 15-JUL-1999, sequence version 2.
DT 07-PEB-2006, entry version 47.
DE Estrogen sulfotransferase, testis isoform (EC 2.8.2.4)
DE (Sulfotransferase, estrogen-prefering).
GN Name=Sult1e1; Synonyms=Ste;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=C57BL/KSJ-DB/DB; TISSUE=Testis;
RX MEDLINE=95269690; PubMed=7750469; DOI=10.1210/en.136.6.2477;
RA Song W.-C., Moore R., McLachlan J.A., Negishi M.;
RT "Molecular characterization of a testis-specific estrogen
RT sulfotransferase and aberrant liver expression in obese and
RT diabetic C57BL/KsJ-db/db mice.";
RL Endocrinology 136:2477-2484(1995).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) IN COMPLEX WITH ADENOSINE
RP 3',5'-BISPHOSPHATE (PAP), AND SEQUENCE REVISION TO 113.
RC STRAIN=C57BL/KSJ-DB/DB; TISSUE=Testis;
RX MEDLINE=98020506; PubMed=9360604;
RA Kakuta Y., Pedersen L.G., Carter C.W., Negishi M., Pedersen L.C.;
RT "Crystal structure of estrogen sulfotransferase.";
RL Nac. Struct. Biol. 4:904-908(1997).
CC -!- FUNCTION: Sulfation of estradiol and estrone. May control the
CC level of the estrogen receptor by sulphylation free estradiol.
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =
CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- TISSUE SPECIFICITY: Testis and at very low level in the placenta.
CC -!- MISCELLANEOUS: Abnormal high expression in liver in obese and
CC diabetic C57BL/KSJ-DB/DB strain mice. Female > male. Normal
CC level in liver.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; S78182; AAB34320.1; -; mRNA.
DR PDB; 1AQY; X-ray; A/B=1-295.
DR PDB; 1AQY; X-ray; A/B=1-295.
DR PDB; 1B06; X-ray; A/B=1-295.
DR IntAct; P49891; -.
DR Ensembl; ENSMUSG00000029272; Mus musculus.
DR MGI; MGI:98431; Sult1e1.

```

```

DR LinkHub; P49891; -.
DR GO; GO:0008210; P:estrogen metabolism; IMP.
DR GO; GO:0007565; P:pregnancy; IMP.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
DR 3D-structure; Lipid-binding; Steroid-binding; Transferase.
KW CHAIN 1 295
FT ESTROGEN SULFOTRANSFERASE, TESTIS
FT isoform.
FT /FTID=PRO_0000085154.
FT PAPS.
FT NP_BIND 48 53
FT NP_BIND 130 138
FT NP_BIND 193 229
FT NP_BIND 257 259
FT ACT_SITE 108 108
FT CONFLICT 113 113
FT HELIX 8 11
FT STRAND 12 15
FT TURN 16 17
FT STRAND 18 21
FT HELIX 22 24
FT TURN 25 27
FT HELIX 28 32
FT TURN 33 33
FT TURN 38 39
FT STRAND 41 45
FT TURN 47 48
FT STRAND 49 50
FT HELIX 51 62
FT TURN 53 64
FT STRAND 75 78
FT STRAND 79 79
FT STRAND 81 81
FT TURN 82 83
FT STRAND 85 86
FT TURN 87 89
FT STRAND 90 91
FT HELIX 92 97
FT TURN 98 98
FT STRAND 101 102
FT STRAND 104 107
FT HELIX 111 113
FT STRAND 114 114
FT HELIX 116 120
FT TURN 121 122
FT STRAND 124 129
FT HELIX 132 145
FT STRAND 146 146
FT TURN 147 148
FT STRAND 149 149
FT STRAND 153 153
FT HELIX 155 164
FT TURN 165 165
FT STRAND 167 167
FT TURN 168 169
FT HELIX 172 182
FT TURN 183 184
FT STRAND 185 185
FT TURN 186 187
FT STRAND 188 192
FT HELIX 193 198
FT TURN 200 210
FT TURN 211 212
FT HELIX 217 226
FT STRAND 227 227
FT STRAND 228 228
FT TURN 229 234
FT STRAND 236 242
FT STRAND 243 243
FT TURN 245 247
FT STRAND 248 248
FT TURN 250 250
FT STRAND 252 253

```

FT STRAND 255 256  
FT STRAND 261 261  
FT HELIX 263 266  
FT TURN 267 267  
FT HELIX 270 284  
FT TURN 285 286  
SQ SEQUENCE 295 AA; 35590 MW; 8E85AB47952BFB1C CRC64;

## Alignment Scores:

Pred. No.: 1.16e-21 Length: 295  
Score: 431.00 Matches: 94  
Percent Similarity: 53.0% Conservative: 57  
Best Local Similarity: 33.0% Mismatches: 116  
Query Match: 9.8% Indels: 18  
DB: 1 Gaps: 3

US-10-768-158-1 (1-2419) x ST1E1\_MOUSE (1-295)

```
QY 60 GAGTTCGAGAGCAAGTACTTCGAGTTCATGGCGTGGCGTCCGCCCTTCTGCGCGCGG 119
Db 7 GluTyrTyrGluValPheGlyGluPheArgGlyValLeuMetAspLysArgPheThrLys 26

QY 120 AAGATGGAGGAGATGCCCAACTTCGCGTGGCGCCAGCGCTGTGGATCGTCACTAC 179
Db 27 TyrTrpGluaspValGluMetPheLeuAlaArgProAspAspLeuValIleAlaThrTyr 46

QY 180 CCCAAGTCCGCGCAGCAGCTTGCTGCAGGAGGTGTCTACTTGGTGAGCCAGCGCGCTGAC 239
Db 47 ProLysSerGlyThrThrTrpIleSerGluValValTyrMetIleTyrLysGluGlyAsp 66

QY 240 CCGATGATCGCGTTCGATGACATCGACGAGCAGTCCGCGTCTCGGATGATCCACAG 299
Db 67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86

QY 300 CCG-----GGCTGCAGCATCATCAGGAAGTACCTCTCCCGGCTCATCAAG 347
Db 87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleValLys 106

QY 348 AGCCACCTGCCTACCGCTTTCTGCGCTCTGACCTCCCAATGGAGACTCCAAGTCTATC 407
Db 107 ThrHisLeuProLysLeuLeuProAlaSerPheThrGluLysAsnCysLysMetIle 126

QY 408 TATATGGCTCGCAACCCCAAGATCTGTGTGTCTTATATATGATTTCCACCGCTCTCTG 467
Db 127 TyrLeuCysArgAsnAlaLysAspValAlaValSerTyrTyrPheLeuLeuMetIle 146

QY 468 CGGACCATGAGTACCGAGGCACCTTTCAGAAATTCGCGGAGGTTCATGATGATGAAG 527
Db 147 ThrSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGlnGlyGln 166

QY 528 CTGGGCTACGGCTCTCTGTTTGGACACGTCGAGGAGTTCTGGGAGCACCGCATGGACTCG 587
Db 167 ValProTyrGlySerTyrTrpAspHisValLysAlaTyrTrpGluLysSerLysAsnSer 186

QY 588 AACGTGCTTTTTCAGATGATGAGACATGATCGGGACCTGGTGATCGATGGTGAGCAG 647
Db 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgArgGluValValLys 206

QY 648 CTGGCCAGATTCCTGGGGGTCTCTGTGACAAAGCCAGCTGGAAGCCCTCAGCGAGCAG 707
Db 207 LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleGlnHis 226

QY 708 TGCCAC-----CAGCTGGTGAGCAGCAG 728
Db 227 ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrMetMetProGluGlu 246

QY 729 TGCTGCAACGCTGAGGCCCTGCC---GTGGCGCGGGAAGAGTTGGCTGTGAAGCAG 785
Db 247 MetMetAsnGlnLysValSerProPheMetArgLysGlyIleIleGlyAspTrpLysAsn 266

QY 786 ATCTTCACCGTCTCCATGAATGAGAGTTTCACCTTGGTGATTAACACAGAAATGGGAAG 845
Db 267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnGlnMetLysAsp 286
```

QY 846 TGTGACCTCACGTTT 860  
Db 287 CysThrValLysPhe 291

## RESULT 56

Q8JZX7\_MOUSE  
ID Q8JZX7\_MOUSE PRELIMINARY; PRT; 295 AA.  
AC Q8JZX7;  
DT 01-OCT-2002, integrated into UniProtKB/TrEMBL.  
DT 01-OCT-2002, sequence version 1.  
DT 07-FEB-2006, entry version 19.

DE Sulfotransferase family 1E, member 1.  
GN Name=Sult1e1;  
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN (1)

## NUCLEOTIDE SEQUENCE.

RP TISSUE=Uterus;

RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Berge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Matra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

## (2)

RP TISSUE=Uterus;

RA Strausberg R.;

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

DR EMBL; BC034891; AAH34891.1; -; mRNA.

DR HSP; P49891; IAAQ.

DR SMR; Q8JZX7; 7-294.

DR Ensembl; ENSMUSG0000029272; Mus musculus.

DR MGI; MGI:98431; Sult1e1.

DR GO; GO:0008210; P:estrogen metabolism; IMP.

DR GO; GO:0007565; P:pregnancy; IMP.

DR InterPro; IPR000863; Sulfotransferase.

DR Pfam; PF00685; Sulfotransfer\_1; 1.

DR ProDom; PD001218; Sulfotransferase; 1.

KW Transferase.

SQ SEQUENCE 295 AA; 35590 MW; 5C2848FBD63E14F4 CRC64;

## Alignment Scores:

Pred. No.: 1.16e-21 Length: 295

Score: 431.00 Matches: 94

Percent Similarity: 53.0% Conservative: 57

Best Local Similarity: 33.0% Mismatches: 116

Query Match: 9.8% Indels: 18

DB: 2 Gaps: 3

```
US-10-768-158-1 (1-2419) x Q8JZX7_MOUSE (1-295)
Qy 60 GAGTTCGAGACAGTACTTCAGTTCATCGCGCTGGCGCTGGCGCCCTTCTGCGCGGG 119
Db 7 GluTyTyTyGluValPheGlyGluPheArgGlyValLeuMetAspLysArgPheThrLys 26
Qy 120 AAGATGGAGGAGATCGCAACTTCCTCGGTGGCGCCAGCGAGTGTGGTACCTAC 179
Db 27 TyTyTyGluValGluMetPheLeuAlaArgProAspAspLeuValIleAlaThrTy 46
Qy 180 CCCAAGTCGCGACACAGCTTCCTGAGGAGTGTCTACTTGGTAGCCAGCGCGCTGAC 239
Db 47 ProLysSerGlyThrTyTyrIleSerGluValValTyMetIleTyTyLysGluGlyAsp 66
Qy 240 CCCGATGAGTCGGTTCGATGACATCGACGAGCTCCGGTCTCGAGTACCCACAG 299
Db 67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyTyLeuGluCysArgAsn 86
Qy 300 CCG-----GGCTCGACATCATCAAGGAAGTCACTCTCCCGCGCTCATCAAG 347
Db 87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleLeuLys 106
Qy 348 AGCCACCTCGCCCTACCGCTTTCTGCGCTCTGACCTCCCAATGGAGACTCCAAAGTCTATC 407
Db 107 ThrHisLeuProProLysLeuLeuProAlaSerPheTyTyGluLysAsnCysLysMetIle 126
Qy 408 TATATGGCTCGAACCCCAAGATCTGTGTGTCTTATTATCATGTTCCACCGCTCTCTG 467
Db 127 TyTyLeuCysArgAsnAlaLysAspValAlaValSerTyTyTyTyTyPheLeuLeuMetIle 146
Qy 468 CGGACCATGAGTACCGAGGACCTTTCAAGAAATCTCCCGAGGTTTATGAATGATAAG 527
Db 147 ThrSerTyTyProAsnProLysSerPheSerGluPheValGluLysPheMetGlnGlyGln 166
Qy 528 CTGGCTACGGCTCTCTGTTTGGACGACGTGAGGAGTCTGGGAGCACCGCATGCTGTCG 587
Db 167 ValProTyTyGlySerTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy 186
Qy 588 AACGTGCTTTTCTCAAGTATGAACATGATCGGACCTGGTGGTGGTGGTGGTGGTGG 647
Db 187 ArgValLeuPheMetPheTyTyGluAspMetLysGluAspIleArgArgGluValValLys 206
Qy 648 CTGGCCAGATCTCTGGGGGTCTCTGTCACAAAGCCAGCTGGAGCCCTGACGGAGCAC 707
Db 207 LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleValGlnHis 226
Qy 708 TGCCAC-----CAGCTGGTGGACCCAG 728
Db 227 ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyTyThrMetMetProGluGlu 246
Qy 729 TGCTGCAACGTGAGCCCTGCCC---GTGGCCGGGGAAGTGGCTGTGGAGGAC 785
Db 247 MetMetAsnGlnLysValSerProPheMetArgLysGlyIleIleGlyAspTrpLysAsn 266
Qy 786 ATCTTCACCGCTCCATGAATGAGAGTTCGACTTGGTGTATAACAGAAAGTGGAAAG 845
Db 267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyTyLysGlnGlnMetLysAsp 286
Qy 846 TGTGACCTCAGCTTT 860
Db 287 CysThrValLysPhe 291
RESULT 57
Q99ND5_RAT
ID Q99ND5_RAT PRELIMINARY; PRT; 295 AA.
AC Q99ND5;
DT 01-JUN-2001, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2001, sequence version 1.
DT 07-FEB-2006, entry version 17.
DE Estrogen sulfotransferase.
GN Name=ste2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Astapova I.I., Yakovenko A.R., Shchelkunova T.A., Chernov B.K.,
RA Sverdlova P.S., Smirnov A.N., Rubtsov P.M.;
RA "Cloning and Preliminary Characterization of the Rat Estrogen
RT Sulfotransferase Gene 5'-Region.";
RL Mol. Biol. 33:447-453 (1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Astapova I.I., Yakovenko A.R., Morozov I.A., Smirnov A.N.,
RA Rubtsov P.M.;
RA "Cloning and Structural Characterization of Two Genes Encoding Rat
RT Estrogen Sulfotransferase.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
TI TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci F., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Kravinsky M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
TI TISSUE=Liver;
RA Director MGC Project;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AJ306223; CAC27405.3; -, Genomic DNA.
DR EMBL; AJ306224; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; AJ306225; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; AJ306226; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; AJ306227; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; AJ306228; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; AJ306229; CAC27405.3; JOINED; Genomic DNA.
DR EMBL; BC088157; AAH88157.1; -, mRNA.
DR HSSP; P49891; 1AQU.
DR SMR; Q99ND5; 7-294.
DR Ensembl; ENSRNOG0000001957; Rattus norvegicus.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 295 AA; 35362 MW; 6A470865BA8F5B0E CRC64;
Alignment Scores: 1.37e-21 Length: 295
Pred. No.: 430.00 Matches: 92
Score: 52.7% Conservative: 62
Percent Similarity:
```

Best Local Similarity: 31.5% Mismatches: 106  
Query Match: 9.8% Indels: 32  
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x Q99ND5\_RAT (1-295)

```
QY 60 GAGTTCGAGCAGCAAGTACTTCAGTTCATGCGTCCGCGTCCGCGCTTCTGCGCGGG 119
DB 7 GluTyrTyrGluValPheGlyAspPheHisGlyValLeuMetAspLysLeuPheThrLys 26
QY 120 AAGATGGAGGAGATGCCCAACTTCGCGTCCGCGCGCAGCGAGTGTGGATGCTGACCTAC 179
DB 27 TyrTrpGluAspValIleThrPheSerAlaIleProAspAspLeuValThrTyr 46
QY 180 CCCAGTCCGCGCAGCAGTTCGCGAGGAGTGTCTACTTGGTGGAGCGCGCTGAC 239
DB 47 ProLysSerGlySerThrTrpIleGlyGluIleValAspMetIleTyrLysGluGlyAsp 66
QY 240 CCCGATGAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGTCTGTGAGTACCCACAG 299
DB 67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86
QY 300 CCG-----GGCTGGACATCATCAAGGAAGTGAACCTCTCCCGCTCATCAAG 347
DB 87 GluAspLeuIleAsnGlyLysGluLeuProAlaSerPheTrpGluLysAsnCysLysIle 106
QY 348 AGCCACCTGCCCTACCGCTTTCTGCGCTCTGACCTCCCAATGGAGACTTCCAAAGTCTATC 407
DB 107 ThrHisLeuProAlaLysLeuLeuProAlaSerPheTrpGluLysAsnCysLysIle 126
QY 408 TATATGGCTCGCAACCCCAAGATCTGGTGTCTTATATCATGTTCCACGCTCTCTG 467
DB 127 TyrLeuCysArgAsnAlaLysAspValValSerTyrTyrPhePheLeuIle 146
QY 468 CGGACCATGAGTACCGAGCACCTTCAAGAACTTCGCGAGGTATTAATGAATGAAG 527
DB 147 LysSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGluGlyGln 166
QY 528 CTGGCTACGGCTCTCGTGTTCGACGACGTGAGGAGTGTCTGGGACGACCGCATGACTCG 587
DB 167 ValProTyrGlySerTyrTrpAspHisValLysSerTyrTrpGluLysSerLysAsnSer 186
QY 588 AACGTGCTTTTCTCAAGTATGAACATGATCGGAGCTGGTGGTGGATGTTGAGGACG 647
DB 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgGluValValLys 206
QY 648 CTGGCCAGATTCCTGGGGTGTCTGTGACAGGCCAGCTGGAGCCCTGACGAGCAC 707
DB 207 LeuIleGluPheLeu-----GluArgAspProLeuAlaGluLeu 219
QY 708 TGCCACCACTGGTGTG-----GACGAGTGTGCAAC 737
DB 220 ValAspLysIleIleGlnHisThrSerPheGlnGluMetLysAsnAsnProCysThrAsn 239
QY 738 GCTGAGGCCCTGCTCC-----GTGGCGCGGGA 764
DB 240 TyrSerMetLeuProGluThrMetIleAspLeuLysValSerProPheMetArgLysGly 259
QY 765 AGATTTGGCTGTGGAAGGACATCTTCACCGTCTCCATGATGACAGACTTTCACCTGTG 824
DB 260 IleValGlyAspTrpArgAsnHisPheProGluAlaLeuArgGluArgPheGluGluHis 279
QY 825 TATAACAGAGATGGGAAGTGTGACCTCAGCTTT 860
DB 280 TyrGlnArgHisMetLysAspCysProValLysPhe 291
RESULT 58
ST1B1 MOUSE
ID ST1B1 MOUSE
AC Q90WG7: O8C301: 0922T0; PRT; 299 AA.
DT 04-JAN-2005, integrated into UniProtKB/Swiss-Prot.
DT 04-JAN-2005, sequence version 2.
DT 07-FEB-2006, entry version 29.
```

DE Sulfotransferase family cytosolic 1B member 1 (EC 2.8.2.-)  
DE (Sulfotransferase 1B1) (DOPA/tyrosine sulfotransferase).  
GN Name=Sulf1b1;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muroidae; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), FUNCTION, TISSUE SPECIFICITY,  
RP AND DEVELOPMENTAL STAGE  
RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;  
RX PubMed=9644246;  
RA Saeki Y., Sakakibara Y., Araki Y., Yanagisawa K., Suiko M.,  
RA Nakajima H., Liu M.-C.;  
RT "Molecular cloning, expression, and characterization of a novel mouse  
RT liver SUL1B1 sulfotransferase.";  
RL J. Biochem. 124:55-64(1998).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).  
RC STRAIN=C57BL/6J; TISSUE=Oviduct;  
RX PubMed=16141072; DOI=10.1126/science.1112014;  
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,  
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,  
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,  
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,  
RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,  
RA Bansal K.P., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,  
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,  
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,  
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gaboldi M.,  
RA Georgi-Hemming P., Gingeras T.R., Gojorori T., Green R.E.,  
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,  
RA Hill D., Huminecki L., Iacono M., Ieko K., Iwama A., Ishikawa T.,  
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,  
RA Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,  
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,  
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,  
RA Matheuda H., Matsuura S., Miki H., Mignone F., Miyake S., Morris K.,  
RA Mortagui-Tabar S., Mulder N., Nakano N., Nakauchi H., Ng P.,  
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,  
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,  
RA Petrovski N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,  
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,  
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,  
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,  
RA Sperling S., Stupka E., Sugura K., Sultana R., Takenaka Y., Taki K.,  
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,  
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,  
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,  
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Ouackenbush J.,  
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,  
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arahawa T.,  
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,  
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Nimomiya N.,  
RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,  
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,  
RA Hayashizaki Y.;  
RT "The transcriptional landscape of the mammalian genome.";  
RL Science 309:1559-1563(2005).  
RN [3]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).  
RC STRAIN=FVB/N; TISSUE=Colon;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S.C., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield A.S., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RL and mouse cDNA sequences";  
 CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -!- FUNCTION: Catalyzes the sulfate conjugation of many hormones,  
 CC neurotransmitters, drugs and xenobiotic compounds. Sulfonation  
 CC increases the water solubility of most compounds, and therefore  
 CC their renal excretion, but it can also result in bioactivation to  
 CC form active metabolites. Sulfates L-DOPA and D-DOPA, tyrosine  
 CC isomers such as DL-m-tyrosine, dopamine and thyroid hormones.  
 CC -!- SUBUNIT: Monomer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q9QMG7-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q9QMG7-2; Sequence=VSP\_012509;  
 CC Note=No experimental confirmation available;  
 CC -!- TISSUE SPECIFICITY: Liver specific.  
 CC -!- DEVELOPMENTAL STAGE: Expression was detected at very low level in  
 CC liver from 1 day-old and then gradually increased to the maximum  
 CC level at 4 weeks old.  
 CC -!- SIMILARITY: Belongs to the sulfotransferase family.  
 CC  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC  
 CC -----  
 CC EMBL; U92076; AAD09249.1; -; mRNA.  
 CC EMBL; AF022894; AAD01746.1; -; mRNA.  
 CC EMBL; AK033498; BAC28321.1; -; mRNA.  
 CC EMBL; AK087595; BAC39939.1; -; mRNA.  
 CC EMBL; BC024361; AA024361.1; -; mRNA.  
 CC PIR; J0186; J0186.  
 CC HSP; P49888; 1HV3.  
 CC SMR; Q9QMG7; 5-236.  
 CC Ensembl; ENSMUSG00000029269; Mus musculus.  
 CC MGI; MGI:2136282; Sult1b1.  
 CC GO; GO:0005829; Cytoplasm; TAS.  
 CC GO; GO:0004062; F-ATP sulfotransferase activity; IDA.  
 CC InterPro; IPR000863; Sulfotransferase.  
 CC Pfam; PF00685; Sulfotransferase; 1.  
 CC ProDom; PD001218; Sulfotransferase; 1.  
 CC Alternative splicing; Lipid metabolism; Steroid metabolism;  
 CC Sulfotransferase.  
 KW CHAIN 1 299 Sulfotransferase family cytosolic 1B  
 FT member 1, 0000085162.  
 FT /FTID=PRO\_0000085162.  
 FT NP\_BIND 48 53 PAPS (By similarity).  
 FT NP\_BIND 131 139 PAPS (By similarity).  
 FT NP\_BIND 194 230 PAPS (By similarity).  
 FT ACT\_SITE 109 109 Proton acceptor (By similarity).  
 FT BINDING 259 259 PAPS (By similarity).  
 FT VARSPIC 299 299 A -> ALAFTNFEIIGSLKPS (in isoform 2).  
 FT /FTID=VSP\_012509.  
 FT CONFLICT 55 55 S -> T (in Ref. 2).  
 FT CONFLICT 137 137 A -> P (in Ref. 1; AAD09249).  
 FT CONFLICT 275 275 E -> K (in Ref. 2).  
 SQ SEQUENCE 299 AA; 34901 MW; BFD0E0909AD8482D CRC64;  
 Alignment Scores:  
 Pred. No.: 1.49e-21 Length: 299  
 Score: 429.50 Matches: 93  
 Percent Similarity: 53.7% Conservative: 51  
 Best Local Similarity: 34.7% Mismatches: 97

Query Match: 9.8% Indels: 27  
 DB: 1 Gaps: 4  
 US-10-768-158-1 (1-2419) x STB1\_MOUSE (1-299)  
 QY 126 GAGGAGATCGCCAACTTCCCGCGCGCCAGCAGCTGGATCGTCACCTACCCCAAG 185  
 DB 29 GluArgileGluGluPheGlnSerThrProGlyAspileValilleThrThrTyProlys 48  
 QY 186 TCGCGCACCACTTGTGTCAGGAGGTCTACTTGTGTGAGCCAGGGCGGTGACCCCGAT 245  
 DB 49 SerGlyThrThrTrpLeuSerGluileValAspMetValLeuAenAspGlyAsnValglu 68  
 QY 246 GAGATCGCGTGTGATGACATCGACGACGCTCCGCTCTGGAGTACCCACACCGCGGC 305  
 DB 69 LysCyslysArgAspValilleThrSerLysValProMetLeuGluLeuSerValProgly 88  
 QY 306 CTG-----GACATCATCAAGAACTGACCTCTCCCGCCCTCATCAAGAGC 350  
 DB 89 IleArgileSerGlyValGluLeuLeuLysThrProSerProAglilleLysThr 108  
 QY 351 CACCTGCGCTACCGCTTTCGCTCTGACCTCCACAAATGGAGACTCCAGGTGATCTAT 410  
 DB 109 HisLeuProIleAspLeuLeuProLysSerPheTrpGluAenLysCysLysMetIleTyr 128  
 QY 411 ATGGCTCGCAACCCCAAGGATCTGGTGTGTCTATTATCAGTTTCCACCGCTCTCTCGG 470  
 DB 129 LeuAlaArgAenGlyLysAspValAlaValSerTyrTyrHisPheAspLeuMetAsnSer 148  
 QY 471 ACCATGAGTACCGGAGGACCTTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTG 530  
 DB 149 IleAsnProLeuProGlyThrTrpGluGluTyrLeuGluLysPheLeuAlaGlyAsnVal 168  
 QY 531 GGCTACGGCTCTCGTTGAGCAGCTGAGAGTCTTGGGAGCACCGCATGGACTCGAAC 590  
 DB 169 AlaTyrGlySerTrpPheAspHisValysSerTrpTrpGluLysArgGluGluHisThr 188  
 QY 591 GTGCTTTTCTCAAGTATGAGACATGCATCGGACCTGGTGACGATGGTGGAGCAGCTG 650  
 DB 189 LeuLeuTyrLeuTyrTyrGluGluLeuLysGlnAsnProLysLysGluileLysLysile 208  
 QY 651 GCCAGATTCTCTGGGGGTCTCTGTGACAAAGCCCGAGCTGGAAGCCCTGACGGAGCAGTGC 710  
 DB 209 AlaSerPheLeuAspLysThrLeuAspGluGluAlaLeuAspArgIleValHisThr 228  
 QY 711 -----CACAGCTGGTGACACCATGCATCGGACCTGGTGACGATGGTGGAGCAGCTG 749  
 DB 229 SerPheGluMetMetLysGluAsnProLeuVal-----AsnTyrThrHisLeu 244  
 QY 750 CCGCTGGGC-----CGGGGAAGAGTGGGCTG 776  
 DB 245 ProThrAlaMetMetAspHisSerLysSerProPheMetArgLysGlyileValGlyAsp 264  
 QY 777 TGAAGGACATCTTCCACCGTCTCCATGAATGAGAAGTTTGAATCTGGTGTATAACAGAG 836  
 DB 265 TrpLysAsnTyrPheThrMetThrGlnThrGluGlnPheAspAlaValTyrLysLysLys 284  
 QY 837 ATGGGAAAGTGTGACCTCACCTTT 860  
 DB 285 MetSerGlyThrThrLeuGluPhe 292  
 RESULT 59  
 Q91W19\_MOUSE PRELIMINARY; PRT; 263 AA.  
 AC Q91W19; integrated into UniProtKB/TrEMBL.  
 DT 01-DEC-2001, sequence version 1.  
 DT 07-FEB-2006, entry version 24.  
 DE Sulfotransferase family 1A, phenol-preferring, member 1.  
 GN Name=Sult1a1;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;



RL Mol. Biol. (Mosk.) 36:635-642(2002).  
 CC -|- FUNCTION: Sulfation of estrone and estradiol. May control the  
 CC level of the estrogen receptor by sulfonylating free estradiol.  
 CC -|- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =  
 CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.  
 CC -|- SUBUNIT: Homodimer (By similarity).  
 CC -|- SUBCELLULAR LOCATION: Cytoplasm.  
 CC -|- TISSUE SPECIFICITY: Liver of young mature males and uterus.  
 CC -|- DEVELOPMENTAL STAGE: Expressed only in the liver of young adult  
 CC animals (100 days old) and is absent in the prepubertal male  
 CC days old), senescent male (800 days old) and female liver.  
 CC -|- INDUCTION: Induced by androgens and suppressed by estrogens. The  
 CC expression is under the influence of pituitary growth hormone and  
 CC thyroid hormone. Is regulated by progesterone in the uterus.  
 CC -|- SIMILARITY: Belongs to the sulfotransferase family.  
 CC  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC  
 CC EMBL; M86758; AA41128.1; -; mRNA.  
 CC EMBL; S76489; AB33441.1; -; mRNA.  
 CC EMBL; AJ131835; CA10515.2; -; Genomic DNA.  
 CC EMBL; AJ298109; CA10515.2; JOINED; Genomic DNA.  
 CC EMBL; AJ298110; CA10515.2; JOINED; Genomic DNA.  
 CC EMBL; AJ298111; CA10515.2; JOINED; Genomic DNA.  
 CC EMBL; AJ298112; CA10515.2; JOINED; Genomic DNA.  
 CC EMBL; AJ298113; CA10515.2; JOINED; Genomic DNA.  
 CC EMBL; AJ298114; CA10515.2; JOINED; Genomic DNA.  
 CC PIR; A41930; A41930.  
 CC HSRP; P49891; LAQU.  
 CC SMR; P49889; 7-294.  
 CC Ensembl; ENSRNOG00000001957; Rattus norvegicus.  
 CC RGD; 3776; See.  
 CC InterPro; IPR000863; Sulfotransferase.  
 CC Pfam; PF00685; Sulfotransferase 1; 1.  
 CC ProDom; PD001218; Sulfotransferase; 1.  
 KW Direct protein sequencing; Lipid-binding; Steroid-binding;  
 KW Transferase.  
 FT CHAIN 1 295 Estrogen sulfotransferase, isoform 3.  
 FT /FTD=PRO\_0000085157.  
 FT NP\_BIND 48 53 PAPS (By similarity).  
 FT NP\_BIND 130 138 PAPS (By similarity).  
 FT NP\_BIND 193 229 PAPS (By similarity).  
 FT NP\_BIND 257 259 PAPS (By similarity).  
 FT ACT\_SITE 108 108 Proton acceptor (By similarity).  
 FT CONFLICT 150 150 P -> Q (in Ref. 2).  
 FT CONFLICT 238 238 T -> I (in Ref. 2).  
 FT CONFLICT 295 295 L -> P (in Ref. 2).  
 SQ SEQUENCE 295 AA; 35416 MW; A77807A21DD2E7EB CRC64;  
 Alignment Scores:  
 Pred. No.: 1.61e-21 Length: 295  
 Score: 429.00 Matches: 92  
 Percent Similarity: 53.3% Conservative: 60  
 Best Local Similarity: 32.3% Mismatches: 115  
 Query Match: 9.7% Indels: 18  
 Db: 1 Gaps: 3  
 US-10-768-158-1 (1-2419) x ST1E3\_RAT (1-295)  
 QY 60 GAGTTCGAGAGCAAGTACTTCGAGTTCCATCGCGTGGCGCTCCGCCCTCTCGCGCGG 119  
 Db 7 GluTyrTyrAspValPheGlyAspPheHisGlyPheLeuMetAspLysArgPheThrLys 26  
 QY 120 AAGATGGAGGAGATCCCAACTTCCCGGTGGCGGCCAGCGAGTGTGGATCGTCACTAC 179  
 Db 27 TyrTrpGluAspValGluThrPheLeuAlaArgProAspAspLeuLeuValThrTyr 46  
 QY 180 CCCAAGTCCGCGACACAGCTTGTGACGAGGTGTCTACTTGGTGGAGCGGCGCTAC 239  
 Db 47 ProLysSerGlySerThrTrpIleValAspMetIleTyrLysGluGlyAsp 66  
 QY 240 CCCGATGAGATCGGGTTGATGAACATCGACGACGAGCTCCCGGTCTCTGGAGTACCCACAG 299

Db 67 ValGluLysCysLysLysGluAspAlaLeuPheAsnArgIleProAspLeuGluCysArgAsn 86  
 QY 300 CCG-----GGCTTGGACATCATCAAGAACTACCTCTCCCGCTCATCAAG 347  
 Db 87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluSerProArgIleValLys 106  
 QY 348 AGCCACCTGGCTACCTCTTCTGCTCTGACCTCCCAATGAGAGACTCCAGGTCTATC 407  
 Db 107 ThrHisLeuProAlaLysLeuLeuProAlaSerPheTrpGluLysAsnCysLysIleIle 126  
 QY 408 TATATGGCTGCAACCAAGGATCTGGTGTGTCTTATTATTCAGTTCCACCGCTCTCTG 467  
 Db 127 TyrLeuCysArgAsnAlaLysAspValValSerTyrTyrTyrPhePheLeuIleMet 146  
 QY 468 CGGACCATGAGTACCGAGGACCTTTCAAGAAATCTGCCGAGGTTTATGAATGATAAG 527  
 Db 147 LysSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGluGln 166  
 QY 528 CTGGCTACCGCTCTGTTTGGACAGTCTGGAGAGTCTGGGAGCACCAGCATGACTCG 587  
 Db 167 ValProTyrGlySerTyrAspHisValLysSerTrpTrpGluLysSerLysAsnSer 186  
 QY 588 AACCTGCTTTTCTCAAGTATGAAGACATCGGACCTGCTGTCACCATGGTGGAGCAG 647  
 Db 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgGluValValLys 206  
 QY 648 CTGCCAGATTCCTGGGGGTCTCTGTGACAAAGCCAGCTGGAGGCCCTGACGAGAC 707  
 Db 207 LeuIleGluPheLeuGluArgAspProSerAlaGluLeuValAspArgIleGlnHis 226  
 QY 708 TGC-----CACGAGCTGGTGGACAGTCTGCAACGCTGAGGCCCTGCC----- 752  
 Db 227 ThrSerPheGlnGluMetLysAsnAsnProCysThrAsnTyrSerMetLeuProGluThr 246  
 QY 753 -----GTGGCGCGGGAAGAGTGGCTGTGGAAGGAC 785  
 Db 247 MetIleAspLeuLysValSerProPheMetArgLysGlyIleValGlyAspTrpLysAsn 266  
 QY 786 ATCTTCACCGTCTCCATGAATGAGAGTTGACTTGGTGTATATAACAGAGATGGGAAG 845  
 Db 267 HisPheProGluAlaLeuArgGluArgPheGluHisTyrGlnGlnMetLysAsp 286  
 QY 846 TGTGACCTCAGCTTT 860  
 Db 287 CysProValLysPhe 291  
 RESULT 61  
 Q3T0S9 BOVIN PRELIMINARY; PRT; 295 AA.  
 AC Q3T0S9;  
 DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.  
 DT 07-MAR-2006, entry version 1.  
 DE Hypothetical protein.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
 OC Pecora; Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=Crossbred x Angus; TISSUE=ileum;  
 RA Moore S., Alexander L., Brownstein M., Guan L., Lobo S., Meng Y.,  
 RA Taniguchi M., Wang Z., Yu J., Prange C., Schreiber K., Shennen C.,  
 RA Wagner L., Bala M., Babazuk S., Barber S., Babakaiff R., Beland J.,  
 RA Chun E., Del Rio L., Gibson S., Hanson R., Kirkpatrick R., Liu J.,  
 RA Matsuo C., Mayo M., Santos R.R., Stott J., Tsai M., Wong D.,  
 RA Siddiqui A., Holt R., Jones S.J., Marra M.A.;  
 RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License



```
CC -----
DR EMBL; BC102274; AA102275.1; -; mRNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Hypothetical protein.
SQ SEQUENCE 295 AA; 34174 MW; A0898231C26B06C6 CRC64;

Alignment Scores:
Pred. No.: 1.61e-21 Length: 295
Score: 429.00 Matches: 95
Percent Similarity: 53.1% Conservative: 57
Best Local Similarity: 33.2% Mismatches: 114
Query Match: 9.7% Indels: 20
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x Q3T0S9_BOVIN (1-295)
Qy 69 AGCAAGTACTCGAGTTCATCGGCTGGCGTCCGCGCTTCTGCGCGGGAAGATGAG 128
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 12 AlalysTyrvallys-----GlyileProleuileLysTyvrPheAlaGluAlaLeugly 29
Qy 129 GAGATCGCAACTTCCCGGTGGCGCCAGCGACGTGTGGATCGTCACCTACCCCAAGTCC 188
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 30 ProleuGluSerPheGluAlaTrpProAspAspLeuLeuileSerThyTrpProlsSer 49
Qy 189 GGCACCACTGCTGCAGAGGTGCTCTACTTGGTGAGCGGGCGCTGACCCCGATGAG 248
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 50 GlyThrThrTrpValSerGluileuAspLeuileTyvrGlnGluGlyAspLeuGluLys 69
Qy 249 ATCGGCTGTAGTGAACATCAGCAGCAGCTCCGCGTCTGGAGTACCCACAGCCG----- 302
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 70 CysGlnArgAlaProValPheLeuArgValProPheLeuGluPheSerAlaProglyVal 89
Qy 303 -----GGCTGGACATCATCAGAACTGACCTCTCCCGCTCATCAAGAGCCACTG 356
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 90 ProThrGlyValGluLeuLysAspThrProAlaProArgLeuLeuLysThrHisLeu 109
Qy 357 CCCTACCGCTTCTGCGCTCTGACCTCCACATGGAGACTCCCAAGTTCATCTATATGCT 416
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 110 ProleuAlaLeuLeuProLysThrLeuLeuAspGlnLysValLysValleTyvrIleAla 129
Qy 417 CGCAACCCCAAGATCTGGTGTCTTATATCATGTTCCACCGCTCTCTCGGACCATG 476
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 130 ArgAsnAlaLysAspValAlaValSerTyvrTyvrHisPheTyvrArgMetAlaLysValHis 149
Qy 477 AGCTACCGAGCACCTTTCAGAAATCTCGCGGAGTTTATGATGATGATGATGCTGCTAC 536
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 150 ProAspProGlyThrTrpAspSerPheLeuGluLysPheMetAlaGlyGluValCysTyvr 169
Qy 537 GGCTCTGTTGTGACACGTGCGAGGTCTGGGAGCACCGCATGGAGCTGCAACGTGCTT 596
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 170 GlySerTrpTyvrGlnHisValGlnGluTrpTrpGluLeuSerHisThiHisProValLeu 189
Qy 597 TTTCTCAGTATGAGACATCATCGGACCTGCTGACGATGGTGAGCAGCTGGCCACA 656
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 190 TyrLeuPheTyvrGluAspIleLysGluAspSerLysArgGluileGlnLysIleLeuGlu 209
Qy 657 TTCCTGGGGGTGCTCTGTGACAGCCCGAGCTGGAGGCCCTG-----ACGAGGAC 707
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 210 PheIleGlyArgSerLeuProGluGluThrValAspHisIleValGlnArgThrsrPhe 229
Qy 708 TGCCACCACTGGTGACAGCTGTCACACGCTGAGGCCCTGCGCGTGGGC----- 758
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 230 LysGluMetLysLysAsnProMetThrAsnTyvrSerThrIleProThrAlaValMetAsp 249
Qy 759 -----CGGGGAAGAGTTGGGCTGTGGAAGGACATCTTCACC 794
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 250 HisSerIleSerAlaPheMetArgLysGlyIleThrGlyAspTrpLysSerThrsrPheThr 269
Qy 795 GTCTCCATGATGAGAAGTTGACTTGGTGTATATAACAGAGATGGGAAGTGTGACCTC 854
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 117 MetLysThrHisLeuProValGlnLeuValProSerPheThrGluGlnAsnCysLys 136
```

```
Db 270 ValAlaGlnAsnGluLeuPheGluAlaHisTyvrAlaLysLysMetAlaGlyCysLysLeu 289
Qy 855 ACGTTTGACTTTTATTTA 872
Dy :::::::::::
Dy 290 ArgPheArgTrpGluLeu 295
RESULT 62
Q90WR6_CHICK
ID Q90WR6_CHICK PRELIMINARY; PRT; 307 AA.
AC Q90WR6;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE Sulfotransferase 1C.
GN Name=SULT1C;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RC TISSUE=Liver;
RX PubMed=15234270; DOI=10.1016/j.abb.2004.05.008;
RA Wilson L.A., Reynolds G.E., Darras V.M., Coughtrie M.W.H.;
RT "cDNA cloning, functional expression, and characterization of chicken
sulfotransferases belonging to the SULT1B and SULT1C families.";
RL Arch. Biochem. Biophys. 428:64-72(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AJ416889; CAC95180.1; -; mRNA.
DR HSP; P50224; 1CJM.
DR Ensembl; ENSGALG0000016805; Gallus gallus.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 307 AA; 36251 MW; 6B76EB00EDA04C9C CRC64;

Alignment Scores:
Pred. No.: 1.92e-21 Length: 307
Score: 428.00 Matches: 87
Percent Similarity: 55.4% Conservative: 61
Best Local Similarity: 32.6% Mismatches: 97
Query Match: 9.7% Indels: 22
DB: 2 Gaps: 2

US-10-768-158-1 (1-2419) x Q90WR6_CHICK (1-307)
Qy 126 GAGAGATCGCACTTCCCGTGGCGCCAGCAGCGTGTGGATCGTCACCTACCCCAAG 185
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 37 AspGlnValTrpAsnPheLysAlaArgProAspAspLeuLeuValAlaThrTyvrAlaLys 56
Qy 186 TCCGGCACCACTGCTGTCAGAGGTGCTACTTGGTGAGCGGCGCTGACCCCGAT 245
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 57 AlaGlyThrThrTrpThrGlnGluIleValAspMetIleGlnAsnGlyAspIleGlu 76
Qy 246 GAGATCGGCTTGATGAACATCGACGAGCTCCCGGTCTCTGGAGTACCCACAGCCG--- 302
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 77 LysCysArgArgAlaSerThrTyvrLysArgHisProPheLeuGluTrpTrpIleProAsp 96
Qy 303 -----GGCTGACATCATCAAGGAAGTACCTCTCCCGCTC 341
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 97 SerSerProLeuGlyTyvrSerGlyLeuLysLeuAlaGluAlaMetProSerProArgThr 116
Qy 342 ATCAAGAGCACCTGCGCTTACCGCTTCTGCGCTCTGACCTCCACATGGAGACTCCCAAG 401
Dy ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Dy 117 MetLysThrHisLeuProValGlnLeuValProSerPheThrGluGlnAsnCysLys 136
```

```

QY 402 GTCATCTATATGGCTCGCAACCCCAAGGACTGTGGTGGTCTTATTATCAGTCTCCACGCG 461
Db 137 IletlelyrvalAlaAgaAsnAlaLysAspAenLeuValSerTyrTyrHisPheHisArg 156
QY 462 TCTCTCGCGACCATGAGCTACCGAGCCACCTTTCAAGAAATTCGCGGAGGTTTATGAAT 521
Db 157 MetAenLysValLeuProAspProGlyThrIleGluGluPheThrGluLysPheMetAsn 176
QY 522 GATAGCTGGGCTACGGCTCTCTGTTGAGCAGCTGCAGGAGTTCTGGGAGCACCGCATG 581
Db 177 GlyGluValLeuTrpGlySerTrpTyrAspHisValLysGlyTrpTrpLysAlaLysAsp 196
QY 582 GACTCGAACGCTTTTCTCAAGTATCAAGACATGCATCGGACCTGGTGACGATGGTG 641
Db 197 LysHisArgIleLeuTyrLeuPheTyrGluAspMetLysGluAenProLysArgGluIle 216
QY 642 GAGCAGCTGGCGACATTCCTCGGGGTCTCTGTCACAAAGGCCACGCTG----- 689
Db 217 GlnLysIleMetLysPheLeuGluLysAspLeuAspGluGluValLeuAsnLysIle 236
QY 690 -----GAAGCCCTGACGGACATGCCACACGCTG 719
Db 237 TyrAenThrSerPheGluIleMetLysAspAenProMetThrAsnTyrThrLysAspPhe 256
QY 720 GTGGACCACTGCTGCAACGCTGAGCCCTGCGCTGGCGCGGGAAGAGTTGGCTGG 779
Db 257 ValGlyValMetAspHisSerValSerProPheMetArgLysGlySerValGlyAspTrp 276
QY 780 AAGCACTCTTCCACGCTCCCATGAATGAGAAGTTTGACTGGTGATATAACAGAAATG 839
Db 277 LysAenTyrPheThrValAlaLeuAsnLysLysPheAspGlnAenTyrLysLysLysMet 296
QY 840 GGAAGTGTGACCTCACCTTT 860
Db 297 AlaAspThrSerLeuValPhe 303

RESULT 63
ST1S3_BRARE
ID ST1S3_BRARE STANDARD; PRT; 301 AA.
AC Q772V2;
DT 31-AUG-2004, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-2003, sequence version 1.
DT 07-FEB-2006, entry version 19.
DE Cytosolic sulfotransferase 3 (EC 2.8.2.-) (SULT1 ST3).
GN Name=sult1st3;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], ACTIVITY, TISSUE SPECIFICITY, AND ENZYME
RP REGULATION.
RX MEDLINE=22630056; PubMed=12745256; DOI=10.1016/S0003-9861(03)00172-3;
RA Sugahara T, Liu C.-C., Carter G., Pai T.G., Liu M.-C.;
RT "cDNA cloning, expression, and functional characterization of a
RT zebrafish SULT1 cytosolic sulfotransferase.";
RL Arch. Biochem. Biophys. 414:67-73(2003).
CC -!- FUNCTION: Catalyzes the conjugation of sulfate to a variety of
CC xenobiotic and endogenous compounds, including dopamine, T3
CC (3,5,3',5'-tetraiodo-L-thyronine), T4 (thyroxine), estrone, DHEA
CC (dehydroepiandrosterone), flavonoids, isoflavonoids and other
CC phenolic compounds.
CC -!- ENZYME REGULATION: Inhibited by Hg(2+), Co(2+), Zn(2+), Cd(2+),
CC Cu(2+) and Pb(2+) ions. Activated slightly by Mn(2+), Ca(2+) and
CC Mg(2+) ions.
CC -!- BIOPHYSICO-CHEMICAL PROPERTIES:
CC pH dependence:
CC Optimum pH is about 5.0 with n-propyl gallate as substrate, and
CC another smaller pH optimum is observed spanning pH 9.5-10.5.
CC Optimum pH is 10.5 with dopamine as substrate;
CC Temperature dependence:
CC Active from 20 to 43 degrees Celsius;

```

```

CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC -----
CC EMBL: AY196985; AAP55637.1; -; mRNA.
CC Ensembl: ENSDARG00000018361; Danio rerio.
CC ZFIN: ZDB-GENE-030804-28; sult1st3.
CC DR GO: 0005737; C:cytoplasm; NAS.
CC DR GO: 0008146; F:sulfotransferase activity; IDA.
CC DR GO: 0006805; P:xenobiotic metabolism; IDA.
CC DR InterPro: IPR000863; Sulfotransferase.
CC DR Pfam: PF00685; Sulfotransferase_1.
CC DR ProDom: PD001218; Sulfotransferase; 1.
CC KW Catecholamine metabolism; Lipid metabolism; Steroid metabolism;
CC TRANSFERASE.
CC CHAIN 1 301 Cytosolic sulfotransferase 3.
CC NP_BIND 53 58 /FTid=PRO_0000085175.
CC NP_BIND 137 145 PAPS (By similarity).
CC NP_BIND 201 237 PAPS (By similarity).
CC NP_BIND 263 265 PAPS (By similarity).
CC ACT_SITE 115 115 Proton acceptor (By similarity).
CC SQ SEQUENCE 301 AA; 35377 MW; 2E4D05DA38EB741 CRC64;
Alignment Scores:
Pred. No.: 2,07e-21 Length: 301
Score: 427.50 Matches: 97
Percent Similarity: 54.1% Conservative: 54
Best Local Similarity: 34.8% Mismatches: 109
Query Match: 9.7% Indels: 19
DB: 1 Gaps: 5
US-10-768-158-1 (1-2419) x ST1S3_BRARE (1-301)
QY 81 GAGTTTCATGGCGTGGCGCTGCGCCCTTCTGCGCGGGAAGATGGAGGATCGCCAAC 140
Db 19 AspPheGluGlyIleSerMetIleHisTyrPheThrAspAenTrpGluLysValLysAsn 38
QY 141 TTCCCGGTGGCGGCCAGCAGCTGTGGATCGTACCTACCCCAAGTCCGGCACCAGCTTG 200
Db 39 PheGlnAlaArgProAspAspIleLeuIleAlaThrTyrProLysAlaGlyThrThrTrp 58
QY 201 CTGCAGGAGGTGCTCTACTTGGTGAGCGAGGCGCTGAC---CCGATGAGATCGGCTTG 257
Db 59 ValSerTyrIleLeuAspLeuLeuTyrPheGlyAsnGluSerProGluArgGlnThrSer 78
QY 258 ATGAACATCGACGAGCAGCTCCCGGTCTGTGAGTACCCACAGCGCGGCTGGACATCATC 317
Db 79 GlnProIleTyrMetArgValProPheLeuGluAlaCysPheGluGlyIleProPheGly 98
QY 318 AAGGAAGCTG-----ACCTCTCCCGCCTCATCAAGAGCCACTGCCCTAC 362
Db 99 ThrGluLeuAlaAspAenLeuProThrSerProArgLeuIleLysThrHisLeuProVal 118
QY 363 CGCTTTCTGCCCTCTGACCTCCACAAATGGAGACTCCCAAGGTCACTATATGGTCGCGAAC 422
Db 119 GlnLeuValProLysSerPheThrGluGlnAsnSerLysValValTyrValAlaArgAsn 138
QY 423 CCCAAGGATCGGGTGGTCTTATTATCAGTTCCACCGCTCTCTGCGGACCATGAGCTAC 482
Db 139 AlaLysAspAenAlaValSerTyrPheHisPheAspArgMetAsnMetGlyGlnProGlu 158
QY 483 CGAGGCACTTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCC 542
Db 159 ProGlyAspTrpAenThrPheLeuGlnLysPheMetGluGlyArgAsnValPheGlyPro 178
QY 543 TGGTTTGAGCAGCTGGCAGGAGTCTCTGGAGCACCAGCATGGAC---TCGAAGTCTTTT 599
Db 179 TrpTyrAspHisValAenGlyTyrTrpLysLysLysGlnThrTyrSerAsnIleLeuTyr 198
QY 600 CTCGAAGTATGAAGACATGCATCGGACCTGGTGACGATGGTGGAGCAGCTGGCCAGATTC 659

```

Db 199 MetPheTyGluAspMetValGluAenThrThyArgGluValGluArgLeuCysserPhe 218  
QY 660 CTGGGGGTCTCTGTGACAAAGCCAGCTGGAAGCCCTGACGGAGCAGCTGCCAC----- 713  
Db 219 LeuGlyLeuSerThrSerAlaAlaGluArgGluArgIleThrLysGlyValGlnPheAsp 238  
QY 714 ---CAGCTGGTGACAGCTGTCMAAGCTGAGGCCCTGCCGTG----- 755  
Db 239 AlaMetLysGlnAenLysMetThrAsnTySerThrIleProValMetAspPheLysIle 258  
QY 756 -----GGCCGGGAGAGTTGGCTGTGAAGACATCTTCAACCGTCTCCATG 803  
Db 259 SerProPheMetArgLysGlyValGlyAspTrpArgAsnHisPheThrValAlaGln 278  
QY 804 AATGAGAAGTTTGACTGTGTATATAACAGAGATGGGAAAGTGTGACCTCACGTTT 860  
Db 279 AenGluGlnPheAspGluValTyLysGlnLysMetLysAenThrThrValLysPhe 297  
RESULT 64  
ST2S2\_BRARE  
ID ST2S2\_BRARE STANDARD; PRT; 301 AA.  
AC Q7ZUS4; Q7TIC7;  
DC 31-AUG-2004, integrated into UniProtKB/Swiss-Prot.  
DT 01-JUN-2003, sequence version 1.  
DT 07-FEB-2006, entry version 2. (EC 2.8.2.-) (SULT1 S2).  
DE Cytosolic sulfotransferase 2  
GN Name=sult1s2;  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [MRNA], ACTIVITY, TISSUE SPECIFICITY, AND ENZYME  
REGULATION.  
RX MEDLINE=22640919; PubMed=12755695;  
RA DOI=10.1046/j.1432-1033.2003.03608.x;  
RA Sugahara T., Liu C.-C., Pai T.G., Collodi P., Suiko M., Sakakibara Y.,  
RA Nishiyama K., Liu M.-C.;  
RT "Sulfation of hydroxychlorobiphenyls. Molecular cloning, expression,  
RT and functional characterization of zebrafish SULT1  
RT sulfotransferases.";  
RL Eur. J. Biochem. 270:2404-2411(2003).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RG NIH - Zebrafish Gene Collection (ZGC) project;  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Catalyzes the conjugation of sulfate to a variety of  
CC xenobiotic and endogenous compounds, including 2-naphthol,  
CC hydroxychlorobiphenyls, T3 (triiodo-L-thyronine), T4 (thyroxine),  
CC estrone and DOPA.  
CC -!- ENZYME REGULATION: Inhibited by Co(2+), Zn(2+), Cd(2+) and Pb(2+)  
CC ions. Inactivated by Hg(2+) and Cu(2+) ions.  
CC -!- BIOPHYSICOCHEMICAL PROPERTIES:  
CC pH dependence:  
CC Optimum pH is 4.75 and 10.5. These two pH optima may correspond  
CC to two distinct conformational states of the enzyme;  
CC Temperature dependence:  
CC Active from 20 to 43 degrees Celsius;  
CC -!- SUBCELLULAR LOCATION: Cytoplasm.  
CC -!- TISSUE SPECIFICITY: Expressed in liver.  
CC -!- SIMILARITY: Belongs to the sulfotransferase family.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; AY181065; AA064984.1; -; mRNA.  
DR EMBL; BC047850; AAH47850.1; -; mRNA.  
DR HSSP; P49888; 1HY3.  
DR Ensembl; ENSDARG00000018361; Danio rerio.  
DR ZFIN; ZDB-GENE-030804-27; sult1s2.  
DR GO; GO:0005737; C:cytoplasm; NAS.

DR GO; GO:0008146; F:sulfotransferase activity; IDA.  
DR GO; GO:0006805; P:xenobiotic metabolism; IDA.  
DR InterPro; IPR000863; Sulfotransferase.  
DR Pfam; PF00685; Sulfotransferase 1; 1.  
DR ProDom; PD001218; Sulfotransferase; 1.  
KW Catecholamine metabolism; Lipid metabolism; Steroid metabolism;  
KW Transferase.  
FT CHAIN 1 301 Cytosolic sulfotransferase 2.  
FT NP\_BIND 53 58 PAPS (By similarity).  
FT NP\_BIND 137 145 PAPS (By similarity).  
FT NP\_BIND 201 237 PAPS (By similarity).  
FT NP\_BIND 263 265 PAPS (By similarity).  
FT ACT\_SITE 115 115 Proton acceptor (By similarity).  
FT CONFLICT 111 111 P -> L (in Ref. 1).  
FT CONFLICT 226 226 T -> A (in Ref. 1).  
FT CONFLICT 239 239 V -> A (in Ref. 1).  
FT CONFLICT 285 285 V -> D (in Ref. 1).  
SQ SEQUENCE 301 AA; 35364 MW; 5943B7C4E33C621C CRC64;  
Alignment Scores:  
Pred. No.: 2.07e-21 Length: 301  
Score: 427.50 Matches: 99  
Percent Similarity: 54.1% Conservative: 52  
Best Local Similarity: 35.5% Mismatches: 109  
Query Match: 9.7% Indels: 19  
DB: 1 Gaps: 5  
US-10-768-158-1 (1-2419) x ST2S2\_BRARE (1-301)  
QY 81 GAGTTCATGGCGTGGCTCGCGCTTCTGCGCGGGAAGATGGAGAGATCCGCCAAC 140  
Db 19 AspPheGluGlyValSerMetThrArgTyPheThrAspAsnTrpGluLysValLysAsn 38  
QY 141 TTCGGGTGGCGCCAGCAGTGTGGATCTCACCTACCCCACTCCGCGGACAGCTTG 200  
Db 39 PheGlnAlaArgProAspIleLeuIleAlaThrTyProLysAlaGlyThrTrp 58  
QY 201 CTGAGGAGGTGCTTACTTGTGAGCGCGCGGTGAC---CCGATGAGATCGGCTTG 257  
Db 59 ValSerTyIleLeuAspLeuLeuTyPheGlyAsnGluSerProGluArgGlnThrSer 78  
QY 258 ATGAACATCGACGACAGCTCCGGTCTCTGGAGTACCCACAGCGCGGCTGACATATC 317  
Db 79 GlnProIleTyMetArgValProPheLeuGluMetCysPheGlnGlyLeuProLeuGly 98  
QY 318 AAGNACTG-----ACCTCTCCCGCTCATCAAGAGTCATCTATATGGCTCGCAAC 362  
Db 99 ThrGluLeuAlaAspThrLeuProThrSerProArgProIleLysThrHisLeuProVal 118  
QY 363 CGCTTTCGCTCTGACCTCCACATGGAGACTCCAAAGTCCATGATGATAGCTGGCTCC 422  
Db 119 GlnLeuValProLysSerPheTrpGluGlnAsnSerLysValValTyValAlaArgAsn 138  
QY 423 CCCAAGGATCTGGTGTCTTATATCATGTTCCACCGCTCTCTGCGGACCATGAGTAC 482  
Db 139 AlalysAspAsnAlaValSerTyPheHisPheAspArgMetAsnMetGlyGlnProGlu 158  
QY 483 CGAGGCACCTTTCAAGAAATCTCGCGGAGGTTTATGAATAGTATGCTGGCTAGGCTCC 542  
Db 159 ProGlyAspTrpAenThrPheLeuGlnLysPheMetAspGlyArgAsnValPheGlyPro 178  
QY 543 TGGTTTGAGCAGCTGCAGAGGAGTTCTGGGAGCAGCCGATGGAC---TCGAAGCTGCTTTT 599  
Db 179 TrpTyArgHisValAsnGlyTyTrpLysLysLysGlnThrTySerAsnIleLeu 198  
QY 600 CTCAGTATGAGACATGCTGCGGACCTGTGTACGATGTGGAGTGTGGAGCTGGCCAGTTC 659  
Db 199 MetPheTyGluAspMetValGluAspThrGlyArgGluValAlaAlaLeuCysSerPhe 218  
QY 660 CTGGGGGTGCTCTGTGACAGGCCAGCTGGAAGCCCTGACGGAGACTGCCACCAGCTG 719  
Db 219 LeuGlyLeuSerThrSerAlaThrGluArgGluArgIleThrLysGlyValGlnPheAsp 238



RESULT 66  
ID Q2XV97\_MACFA PRELIMINARY; PRT; 294 AA.  
AC Q2XV97;  
DT 20-DEC-2005, integrated into UniProtKB/TrEMBL.  
DT 20-DEC-2005, sequence version 1.  
DT 07-FEB-2006, entry version 3.  
DE Sulfotransferase family 1b, estrogen-preferring, member 1.  
GN Name=SULT1E1;  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
OC Cercopithecoidea; Cercopithecinae; Macaca.  
ON NCBI\_TaxID=9541;  
RX NCBI\_TaxID=9541;  
RN NUCLEOTIDE SEQUENCE.  
RP Liu H., Larbie F., Luu-The V.;  
RL Submitted (SEP-2005) to the EMBL/GenBank/DBJ databases.  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC  
DR EMBL: DQ228171; ABB76812.1; -: mRNA.  
DR GO: GO:0016740; F:transferase activity; IEA.  
KW Transferase.  
SQ SEQUENCE 294 AA; 35159 MW; B9997AFDBCC353C0 CRC64;  
  
Alignment Scores:  
Pred. No.: 4,28e-21 Length: 294  
Score: 423.00 Matches: 91  
Percent Similarity: 52.3% Conservative: 58  
Best Local Similarity: 31.9% Mismatches: 118  
Query Match: 9.6% Indels: 18  
DB: Gaps: 3  
  
US-10-768-158-1 (1-2419) x Q2XV97\_MACFA (1-294)  
QY 60 GAGTTCGAGACGAAGTACTTCGAGTTCCTCCGCGTGGCGTCCGCGCTTCGCGCGGG 119  
DB 6 AspTyrTyrGluAsnPheGluGluLeuHisGluValLeuMetTyrLysAspPheValLys 25  
QY 120 AAGATGAGGAGATCGCAACTTCGCGTGGCGCCGAGGAGTGGTGGATCGTCACTAC 179  
DB 120 AAGATGAGGAGATCGCAACTTCGCGTGGCGCCGAGGAGTGGTGGATCGTCACTAC 179  
QY 26 TyrTrpAsnAspValGluThrPheGlnAlaArgProAspAspLeuValIleAlaThrTyr 45  
DB 26 TyrTrpAsnAspValGluThrPheGlnAlaArgProAspAspLeuValIleAlaThrTyr 45  
QY 180 CCAAGTCCGCGACCAAGTTCGAGGAGTGGTCTACTTGGTGGAGCCAGGCGCTGAC 239  
DB 46 ProLysSerGlyThrTyrTrpValSerGluIleAlaTyrMetIleTyrLysGluGlyAsp 65  
QY 240 CCGCATGAGATCGCTTGATGAACATCGACGAGCTCCGCGTCTCGAGTACCCACAG 299  
DB 66 ValGluLysCysLysGluAspValIlePheAsnArgIleProPheLeuGluCysArgLys 85  
QY 300 CCG-----GCGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAG 347  
DB 86 GluAspLeuMetAsnGlyValLysGlnLeuAspGluMetAsnSerProArgIleValLys 105  
QY 348 AGCCACTGCGCTACCGCTTTCGCTCTGACCTCCACAAATGGAGACTCCAAAGTCAATC 407  
DB 106 ThrHisLeuProGluLeuLeuProAlaSerPheTrpGluLysAsnCysLysIleIle 125  
QY 408 TATATGCTCGCAACCCCAAGGATGGTGGTCTTATTATTCAGTTCCACGCTCTCTG 467  
DB 126 TyrLeuCysArgAsnAlaLysAspValAlaValSerPheTyrTyrPhePheLeuMetVal 145  
QY 468 CGGACCATGAGTACCGAGGACCTTTCAAGAAATCTCCGCGAGGTTTATGAATGAAG 527  
DB 146 AlaGlyHisProAsnProGlySerPheProGluPheValGluLysPheMetGlnGlyGln 165  
QY 528 CTGGGCTACGGCTCTCGTTTGAAGTCAGGTCAGGATTTCTGGGAGCACCGCATGCACTCG 587  
DB 166 ValProTyrGlySerTyrTrpTyrLysHisValLysSerTrpTrpGluLysGluLysSerPro 185

RESULT 66  
 Q2XV97\_MACFA PRELIMINARY; PRT; 294 AA.  
 AC Q2XV97;  
 DT 20-DEC-2005, integrated into UniProtKB/TrEMBL.  
 DT 20-DEC-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 3.  
 DE Sulfotransferase family 1b, estrogen-preferring, member 1.  
 GN Name=SULT1B1;  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 OC Cercopithecoidea; Cercopithecinae; Macaca.  
 ON NCBI\_TaxID=9541;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Liu H., Larbie F., Luu-The V.;  
 RL Submitted (SEP-2005) to the EMBL/GenBank/DDBJ databases.  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC  
 DR EMBL: DQ228171; ABB76812.1; -: mRNA.  
 DR GO: GO:0016740; F:transferase activity; IEA.  
 KW Transferase.  
 SQ SEQUENCE 294 AA; 35159 MW; B9997AFDBCC353C0 CRC64;  
 Alignment Scores:  
 Pred. No.: 4,28e-21 Length: 294  
 Score: 423.00 Matches: 91  
 Percent Similarity: 52.3% Conservative: 58  
 Best Local Similarity: 31.9% Mismatches: 118  
 Query Match: 9.6% Indels: 18  
 DB: Gaps: 3  
 US-10-768-158-1 (1-2419) x Q2XV97\_MACFA (1-294)  
 QY 60 GAGTTCGAGACGAAGTACTTCGAGTTCCTCCGCGTGGCGTCCGCCCTTCTCGCGCGG 119  
 DB 6 AspTyrTyrGluAsnPheGluGluLeuHisGluValLeuMetTyrLysAspPheValLys 25  
 QY 120 AAGATGAGAGAGATCGCAACTTCCTCCGCGTGGCGCCAGGAGTGGTGGATCGTCACTAC 179  
 DB 26 TyrTrpAsnAspValGluThrPheGlnAlaArgProAspAspLeuValIleAlaThrTyr 45  
 QY 180 CCAAGTCCGCGACCAAGTCTGTCGAGGAGTGTCTACTTGGTGGACCCAGGCGCTGAC 239  
 DB 46 ProLysSerGlyThrTrpValSerGluIleAlaTyrMetIleTyrLysGluGlyAsp 65  
 QY 240 CCGCATGAGATCGCTTGATGAACATCGACGACGCTCCCGTCTCTGGAGTACCCACAG 299  
 DB 66 ValGluLysCysLysGluAspValIlePheAsnArgIleProPheLeuGluCysArgLys 85  
 QY 300 CCG-----GCGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAG 347  
 DB 86 GluAspLeuMetAsnGlyValLysGlnLeuAspGluMetAsnSerProArgIleValLys 105  
 QY 348 AGCCACTGCGCTACCGCTTCTGCGCTCTGACCTCCAAATGGAGACTCCAAAGTCAATC 407  
 DB 106 ThrHisLeuProGluLeuLeuProAlaSerPheTrpGluLysAsnCysLysIle 125  
 QY 408 TATATGCTCGCAACCCCAAGGATGTTGGTGTCTTATTATCAGTTCCACGCTCTCTG 467  
 DB 126 TyrLeuCysArgAsnAlaLysAspValAlaValSerPheTyrTrpPheLeuMetVal 145  
 QY 468 CGGACCATGAGTACCGAGGACCTTTCAAGAAATCTCCGCGAGGTTTATGAATGAAG 527  
 DB 146 AlaGlyHisProAsnProGlySerPheProGluPheValGluLysPheMetGlnGlyGln 165  
 QY 528 CTGGGCTACGGCTCTGTTTGAAGTCAGGTCAGGATTTCTGGGAGCACCGCATGCACTCG 587  
 DB 166 ValProTyrGlySerTrpTyrLysHisValLysSerTrpTrpGluLysGluLysSerPro 185

[illegible]

```
QY 300 CCGGCCTGGACATCATCAAGAACTG---ACCTCTCCCGCTCATCAAGACCCACCTG 356
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
100 SerGlyThrGluMetAlaAspAenLeuProThrSerProArgLeuIleLysThrHisLeu 119
QY 357 CCTACCGCTTCTCGCTCTGACCTCCACATGAGACTCAAGGTCTATCTATGCT 416
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
120 ProValGlnLeuValProLysSerPheTrpGluGlnAenSerArgValValTrpValAla 139
QY 417 CCACACCCCAAGGATCTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATG 476
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
140 ArgAsnAlaLysAspAsnAlaValSerTyPheHisPheAspArgMetAsnMetValGln 159
QY 477 AGCTACCGAGGACCTTTCAAGAAATTCGCGGAGTTTATGAATGATAAGCTGGCTAC 536
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
160 ProAspProGlyAspTrpAspSerTyPheLeuAspLysPheMetGlnGlyGlnAsnValPhe 179
QY 537 GGCTCTCTGTTGACACGTCGAGAGTCTGGAG---CACCGCATGGACTCGAACGTG 593
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
180 GlySerTrpPheAspHisValSerGlyTrpTrpGlnLysLysAspArgMetAsnMet 199
QY 594 CTTTCTTCAGTATGAACATGATCGGACCTGGTGACGATGGTGAGCAGCTGGCC 653
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
200 LeuTyMetPhePheGluAspLeuSerGluAspThrGlyArgGluValAsnArgLeuCys 219
QY 654 AGATTCTGGGGTCTCTGTGACAAAGCCAGCTGGAAGCCCTGACGAGCAGCTGCCAC 713
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
220 SerPheLeuGlyLeuSerThrSerValGlnGluLysGluLysIleThrLysGlyValGln 239
QY 714 -----CAGCTGGTGACGAGCTGCAACGCTGAGGCCCTGCC----- 752
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
240 PheAspAlaMetLysGlnAsnThrLeuIleAsnHisValThrIleProPheLeuAspCys 259
QY 753 -----GTGGCGCGGGAAGAGTGGCTGTGAAGGACATCTTACCGCTC 797
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
260 LysIleSerProPheMetArgLysGlyLysValGlyAspTrpLysSerHisPheThrVal 279
QY 798 TCCATGAATGAGAGTTTGTGCTGATATAACAGAGATGGGAAAGTGTGACCTCAG 857
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
280 AlaGlnAsnGluArgPheAspGluValTyLysGlnLysMetLysAsnSerGlyValThr 299
QY 858 TTT 860
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
300 Phe 300

RESULT 69
ST1E1_RAT
ID ST1E1_RAT STANDARD; PRT; 295 AA.
AC P52844;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-1996, sequence version 1.
DE Estrogen sulfotransferase, isoform 1 (EC 2.8.2.4) (EST-1)
DE (Sulfotransferase, estrogen-prefering) (Estrone sulfotransferase).
GN Name=St1e1; Synonyms=Ste, Stel;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=Fischer 344; TISSUE=Liver;
RA MEDLINE=96305357; PubMed=8688469; DOI=10.1016/0167-4781(96)00065-6;
RT "Structural relationships among members of the mammalian
RT sulfotransferase gene family."
RL Biochim. Biophys. Acta 1307:331-338(1996).
CC -!- FUNCTION: Sulfation of estrone and estradiol. May control the
CC level of the estrogen receptor by sulfonylating free estradiol (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =
CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.
```

```
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL: U50204; AA07680.1; -; mRNA.
DR HSSP: P49891; 1B06.
DR SMR: P52844; 7-294.
DR Ensembl: ENSRNOG0000001957; Rattus norvegicus.
DR RGD: 3776; Ste.
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransferase_1; 1.
DR ProDom: PD001218; Sulfotransferase; 1.
KW Lipid-binding; Steroid-binding; Transferase.
FT CHAIN 1 295 Estrogen sulfotransferase, isoform 1.
FT NP_BIND 48 53 /FTID=PRO_0000085155.
FT NP_BIND 130 138 PAPS (By similarity).
FT NP_BIND 193 229 PAPS (By similarity).
FT NP_BIND 257 259 PAPS (By similarity).
FT ACT_SITE 108 108 Proton acceptor (By similarity).
SQ SEQUENCE 295 AA; 35509 MW; 696A12FDA923A12E CRC64;

Alignment Scores:
Pred. No.: 5.94e-21 Length: 295
Score: 421.00 Matches: 91
Percent Similarity: 53.0% Conservative: 60
Best Local Similarity: 31.9% Mismatches: 116
Query Match: 9.6% Indels: 18
DB: 1 Gaps: 3

US-10-768-158-1 (1-2419) x ST1E1_RAT (1-295)
QY 60 GAGTTCGAGACAGTACTTCGAGTTCATCGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 119
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7 GluTyTrpGluValPheGlyAspPheHisGlyPheLeuMetAspLysArgPheThrLys 26
QY 120 AAGATGAGGAGATCGCAACTTCGCGGTGGCGGCCAGCGCTGTGTGATCGTCACTTAC 179
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
27 TyTrpGluAspLleGluThrPheLeuAlaArgProAspAspLeuLeuIleValThrTy 46
QY 180 CCCAAGTCGGCACAGCTTCTGCAGAGAGTGTCTACTTGGTGAGCCAGGGCGCTCAC 239
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
47 ProLysSerGlySerThrTrpIleSerGluIleValAspMetIleTyTrpLysGluGlyAsp 66
QY 240 CCCGATGAGTCGGCTTGATGAACATCGACGACGAGCTCCCGCTCTGGAGTACCCACAG 299
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67 ValGluLysCysLysLysGluAspAlaLeuPheAsnArgIleProAspLeuGluCysArgAsn 86
QY 300 CCG-----GGCTGGACATCATCAAGGAAGTACCTCTCCCGCTCATCAAG 347
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
87 GluAspLeuIleAsnGlyLysGlnLeuLysGluLysGluSerProArgIleValLys 106
QY 348 AGCCACCTGCGCTACCGCTTTCTCGCTCTGCACCTCCCAATGGAGACTCCAAGGTCTATC 407
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
107 ThrHisLeuProAlaLysLeuLeuProAlaSerPheTrpGluLysAsnCysLysIleIle 126
QY 408 TATATGGCTCGCAACCCCAAGGATCTGGTGTGTCTTATTATCATGATTCACCGCTCTCTG 467
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
127 TyrLeuCysArgAsnAlaLysAspValValSerTyTrpTyTrpPhePheLeuIleMet 146
QY 468 CGGACCATGAGTACCAGGAGGACCTTTCAAGAAATTCGCCGAGGTTTATGAATGATTAAG 527
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
147 LysSerTyProAsnProLysSerPheSerGluPheValGluLysPheMetGluGlyGln 166
QY 528 CTGGGCTACGGCTCTCTGGTTTGGACACGTGCAGAGTTCTGGGAGCACCCGATGGACTCG 587
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
167 ValProTyGlySerTyTrpAspHisValLysSerTrpTrpGluLysSerLysAsnSer 186
QY 588 AACGTGCTTTTCTCAAGTATGAAGACATGCATCGGACCTGGTGTGACGATGGTGGACGAG 647
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgArgGluValValLys 206
QY 648 CTGGCCCATTCCTGGGGGTGCTGTGACAAAGCCAGCTGGAAGCCCTGACGAGCAC 707
Db 207 LeuIleGluPheLeuGluArgAspProSerAlaGluLeuValAspArgIleIleGlnHis 226
QY 708 TGC-----CACCAGCTGTGCACGAGCTGCTGCAACGCTGAGGCCCTGCC----- 752
Db 227 ThrSerPheGlnGluMetLysAsnAsnProCysThrAsnTyrSerMetLeuProGluThr 246
QY 753 -----GTGGCCCGGGGAAGAGTGGCTGTGGAAGAC 785
Db 247 MetIleAspLeuLysValSerProPheMetArgLysGlyIleValGlyAspTrpArgAsn 266
QY 786 ATCTTCACCGCTCCATGAATGAGAGTTGACTTGGTGTATTAACACAAAGATGGGAAG 845
Db 267 HisPheProGluAlaLeuArgGluArgPheGluGluHisTyrGlnArgHisMetLysAsp 286
QY 846 TGTGACCTCACGTTT 860
Db 287 CysProValLysPhe 291

RESULT 70
ID Q95JCS_ORNAN PRELIMINARY; PRT; 295 AA.
AC Q95JCS57
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 15.
DE Sulfotransferase SULF1A.
OS Ornithorhynchus anatinus (Duckbill platypus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.
OX NCBI_TaxID=9258;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RA Bolton-Grob R.M., McManus M.E.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AV044182; AAK72405.1; -; mRNA.
DR HSSP; P50224; 1CJM.
DR SMR; Q95JCS5; 8-235.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 295 AA; 34189 MW; 8528DA3ECC8833B7 CRC64;

Alignment Scores:
Pred. No.: 7e-21 Length: 295
Score: 420.00 Matches: 95
Percent Similarity: 54.2% Conservative: 61
Best Local Similarity: 33.0% Mismatches: 113
Query Match: 9.5% Indels: 19
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x Q95JCS_ORNAN (1-295)
QY 53 CCGGGGAGTTTCGAGAGCAAGTACTTCGAGTTCATGCGGTGCGGCTGCCCTTCGT 112
Db 5 ProAspThrSerArgProProValArgVal-AsnGlyIleProLeuIleLysTyrPh 24
QY 113 CCGCGGAGATGGAGGAGATCGCAACTCCCGGTGCGGCCGACGAGCTGTGATCGT 172
Db 24 eAlaGluAsnLeuGlyAlaLeuGluSerLeuArgValGlnProSerAspLeuValSe 44
QY 173 CACCTACCCCAAGTCCGGCACCAGCTTGTGTCAGGAGGTGGTCTACTTGGTGAGCCAGG 232
```

```
Db 44 rThrTyrProLysSerGlyThrThrTrpValSerGluIleLeuAspMetIleTyrGlnG 64
QY 233 CGCTGACCCCGATGATCGCTTGATGAACATCGACGAGCAGCTCCGGTCTCGGAGTA 292
Db 64 yGlyAspLeuGluLysCysGlnArgAlaProValPheLeuArgValProPheLeuGluPh 84
QY 293 CCCACAGCG-----GGCCTGGACATCATCAAGGAACCTGACCTCTCCCGCCT 340
Db 84 eSerIleProGlyMetProSerGlyMetGluThrLeuLysAspThrProSerProAArgLe 104
QY 341 CATCAAGACCCCTGCTCCCTTCTGCTCCCTCTGACCTCCACAAATGGAGACTCCAA 400
Db 104 uLeuLysThrHisLeuProLeuAlaLeuLeuProLysValLeuLeuAspGlnLysVal 124
QY 401 GGTCACTATATGCTCGCAACCCCAAGGATCTGGTGTCTTATTTATCATGATTCACCG 460
Db 124 sValIleTyrIleAlaArgAsnAlaLysAspValAlaValSerTyrTyrHisPheTyrAr 144
QY 461 CTCTCTCGGACCATGAGCTACCGAGGCACCTTCAAGAAATTCGCGGAGGTTTATGAA 520
Db 144 gMetAlaLysValHisProAspProGlyThrTrpGluThrPheLeuGluAlaPheLysAl 164
QY 521 TGATAAGCTGGCTACCGCTCTCTGTTGACACGTCGAGAGTTCCTGGGAGCACCGCAT 580
Db 164 aGlyGlnValCysTyrGlySerTrpTyrGlnHisValGlnAspTrpTrpGluLeuArgLy 184
QY 581 GGACTCGNACGCTGCTTTCTCAAGATGAGACATGCATCGGAGCCTGGTGAGCATGGT 640
Db 184 sGlnGlnProValLeuTyrLeuPheGluAspIleLysGluAspProLysArgGluIl 204
QY 641 GGACGAGCTGCCAGATTCTCTGGGGGTGCTCTGTGACAAGGCCCGCTGGAAGCCCTG-- 698
Db 204 eArgLysIleMetGluPheIleAspGlnProValSerGluGluIleLeuAspArgIleVa 224
QY 699 -----ACGAGCAGCTGCCACAGCTGGTGGACAGCTGCTGCAACGCTGAGGCCCTGCC 751
Db 224 lGlnGluThrSerPheLysLysMetLysGlnAsnProMetThrAsnTyrSerSerValPr 244
QY 752 C-----GTGGCCCGGGAAGAGTTCGGCTGTG 778
Db 244 oSerHisIleMetAspHisGluValSerProPheMetArgLysGlyThrProGlyAspTr 264
QY 779 GAAGGACATCTTACCCTCTCCATGAATGAGAAGTTTGACTTGTGTATTAACAGAAGAT 838
Db 264 pLysAsnHisPheThrValAlaGlnAsnGluLeuPheAspValAspTyrAlaGluLysMe 284
QY 839 GGGAAAGTGTGACCTCACGTTT 860
Db 284 tAlaGlySerGluLeuHisPhe 291

RESULT 71
QYHYKO_STRPU PRELIMINARY; PRT; 285 AA.
AC Q3HYKO;
DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 08-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Sulfotransferase.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Heyland A., Moroz L.L., Price D.A., Bodnarova M.;
RT "Thyroid hormone metabolism and thyroid peroxidase function in non-
RT chordate animals.";
RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
```



```
CC -----
DR EMBL; DQ176319; ABA41638.1; -; mRNA.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 285 AA; 33378 MW; B27DCDDBI7F06518 CRC64;

Alignment Scores:
Pred. No.: 7 52e-21 Length: 285
Score: 419.50 Matches: 94
Percent Similarity: 53.7% Conservative: 59
Best Local Similarity: 33.0% Mismatches: 105
Query Match: 9.5% Indels: 27
DB: Gaps: 5

US-10-768-158-1 (1-2419) x Q3HYKO_STRPU (1-285)
QY 99 CTGCGCCCTTCTGCGCGGGAAG---ATGGAGGAGATCGCAACTTCCCGTGGCGGCC 155
DB 1 MetProLysPheThrProLysArgPheLeuAspLeuLeuAsnPheGluValLysGly 20
QY 156 AGCGAGGTGGATCGTCACCTACCCCAAGTCCCGCACCAGCTTCTCGCAGGAGTGTC 215
DB 21 AspAspThrTyrLeuIleThrTrpProLysSerGlyThrTrpMetGlnAsnIleLeu 40
QY 216 TACTTGGTGAGCGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGACGACGAG 275
DB 41 ThrLeuIlePheAlaLysGlyAspMetAspAlaValArgGluLysHisLeuPheLysArg 60
QY 276 CTCCGCGCTCTGGAGTACCCACAG-----CCGGCC 305
DB 61 ValProPheLeuGluMetProLysGlyPheAspTyrLysLysAlaGluAspAspThrGly 80
QY 306 CTG---GACATCATCAAGAACTGACCTCTCCCGCCTCATCAAGAGCACCTGCGCTAC 362
DB 81 LeuTyrGluIleValArgAsnValProSerProArgLeuLeuLysThrGlnLeuProPro 100
QY 363 CGCTTTCGCTCTGACCTCCACAAATGGAGACTCCAAGGTCATATATGGCTGCGCAAC 422
DB 101 ProPheLeuProThrGlnIleHisGluLysProLysIleValTyrValAlaArgAsn 120
QY 423 CCCAAGGATCTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGCTAC 482
DB 121 ProLysAspAlaAlaValSerTyrPheHisPheCysAsnValSerProAsnLeuProGln 140
QY 483 CGAGGCCTTCAAGATTCTGCGGAGGTTTATGATGATAGTACCTGGGCTACGCTCC 542
DB 141 TyrArgAspTrpAsnAspPhePheIleAspPheCysAsnAspSerIleProArgGlySer 160
QY 543 TGGTTTGACGAGTCAGAGTTCTGGGACACCGCATGCGACTCGAAGTCTTTTCTC 602
DB 161 TrpPheGluAsnValLeuTyrTrpAsnLysArgHisGluSerAsnValLeuPheIle 180
QY 603 AAGTATGAAGCATGCATCGGACCTGGTGACGATGGAGCGACCGCATGCGACGATTCCTG 662
DB 181 ThrTyrGluGluMetLysGlnAspLeuArgGlySerValValArgValCysAspPheLeu 200
QY 663 GGGGTGTCCTGTGACAAGCCCGAGCTGGAAGCCCTGACGGAGCACTGC-----CAC 713
DB 201 GlyLysGluLeuSerAspAspIleIleAspValIleThrGluAsnSerThrPheAsnAla 220
QY 714 CAGCTGGTGGACGAGTGTGCAACCTGAGGCCCTGCGCCGTG----- 755
DB 221 MetLysLysAspProThrAlaAsnProAspSerLeuLeuValPheLysGluAlaLys 240
QY 756 -----GGCCGGGAAGAGTTGGCTGTGGAGGACATCTTACCCGTC 797
DB 241 GlnLysArgSerPheLeuArgLysGlyGluValGlyAspTrpLysAsnHisPheThrVal 260
QY 798 TCCATGAATGAGAAGTTTGTGTTGTTATAAACAGAAGATGGGAAAGTGTGACCTCAG 857
```

---

```
Db 261 AlaGlnAsnIleIlePheAspAspLeuTyrArgGlyLysThrAsnGlySerGlyMetAsn 280
QY 858 TTTGACTTTTATTTA 872
Db 281 PheThrPheGluLeu 285

RESULT 72
SUAC_RAT
ID SUAC_RAT STANDARD; PRT; 304 AA.
AC P50237;
DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-1996, sequence version 1.
DE 07-FEB-2006, entry version 38.
DE N-hydroxyarylamine sulfotransferase (EC 2.8.2.-) (HAST-I).
GN Name=St1c1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Murioidea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND PROTEIN SEQUENCE OF 10-20 AND 28-38.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=94043328; PubMed=8227031;
RA Nagata K., Ozawa S., Miyata M., Shimada M., Gong D.-W., Yamazoe Y.,
RA Kato R.;
RT "Isolation and expression of a cDNA encoding a male-specific rat
RT sulfotransferase that catalyzes activation of N-hydroxy-2-
RT acetylaminofluorene.";
RL J. Biol. Chem. 268:24720-24725 (1993).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA].
RX MEDLINE=94306556; PubMed=8033246; DOI=10.1016/0009-2797(94)90057-4;
RA Yamazoe Y., Nagata K., Ozawa S., Kato R.;
RT "Structural similarity and diversity of sulfotransferases.";
Chem. Biol. Interact. 92:107-117 (1994).
CC -!- FUNCTION: Sulfation of phenols and bioactivation of N-
CC hydroxyarylamines. Is responsible for the formation of N-hydroxy-
CC 2-acetylaminofluorene, a reactive metabolite which exhibits
CC toxicity by binding to DNA, RNA and protein.
CC -!- SUBCELLULAR LOCATION: Cytoplasm (Potential).
CC -!- TISSUE SPECIFICITY: Liver. Male >> Female.
CC -!- DEVELOPMENTAL STAGE: Male specific. Maximum at 9 weeks and
CC maintained in 9-month-old rats. Can be detected at low level in
CC females up to 9-week-old rats but then decreases to undetectable
CC level.
CC -!- INDUCTION: Induced by estrogens and suppressed by androgens.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
-----
Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL; L22339; ABA42181.1; -; mRNA.
PIR; A49098; A49098.
HSSP; P50224; 1CJM.
LinkHub; P50237; -.
GO; GO:0005829; C:cytosol; NAS.
GO; GO:0004062; F:aryl sulfotransferase activity; IDA.
InterPro; IPR000863; Sulfotransferase.
Pfam; PF00685; Sulfotransferase; 1.
ProDom; PD001218; Sulfotransferase; 1.
Direct protein sequencing; Transferase.
KW CHAIN 1 304 N-hydroxyarylamine sulfotransferase.
FT NP_BIND 56 61 /FTid=PRO_000085138.
FT NP_BIND 139 147 PAPS (By similarity).
FT NP_BIND 202 238 PAPS (By similarity).
FT NP_BIND 266 268 PAPS (By similarity).
FT ACT_SITE 117 117 Proton acceptor (By similarity).
SQ SEQUENCE 304 AA; 35764 MW; C772B2EA7BD74198 CRC64;
```



```
QY 306 CTGGACATCATCAAGAACTGACCTCTCCCGCCCTCATCAAGAGCCACTGCCCTACCGC 365
Db 102 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysHisLeuProValHis 121
QY 366 TTTCTGCCCTCTGACCTCCACAAATGGAGACTCCAAGGTTCATATATATGGCTCGCAACCCC 425
Db 122 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAsnAla 141
QY 426 AAGGATCTGGTGGTGTCTTATTATATCAGTTTCACCGCTCTCTCGGACCATGAGTACCGA 485
Db 142 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuLeuProAspPro 161
QY 486 GGCACCTTTCAGAAATTCGCGGAGGTTTATGAATGATAGTAACTGGCTACGGCTCCCTGG 545
Db 162 GlyThrLeuGlyGlyIleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 181
QY 546 TTTGAGCAGCTGCAGAGTTCTGGGAGCACCGCATCGGACTCGAAAGCTGTTTTTCTCAAG 605
Db 182 TyrAspHisValLysGlyTrpTrpAspValLysAspGlnHisArgIleLeuTyrLeuPhe 201
QY 606 TATGAAGACATGCATCGGAGCTGTGACGATGTGTGGAGCAGCTGGCAGATTCCTGGGG 665
Db 202 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 221
QY 666 GTGTCTCTGACAAAGCCAGCTGAAGCCCTGACGGAGCAGTGGCCACCAGCTGGTG--- 722
Db 222 LysAspIleSerGluGluValLeuAsnLysIleIleTyrHisThrSerPheAspValMet 241
QY 723 -----GACCAGTGTGCAACGCTGAGGCCCTGCCCC----- 752
Db 242 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 261
QY 753 -----GTGGCGCGGAGAGTGTGGCTGTGGAAGGAGCATCTTCACCGTCTCCATG 803
Db 262 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 281
QY 804 AATGAGAAGTTGACTTGGTGATATAACAGAGATGGGAAGTGTGACCTACAGTTT 860
Db 282 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 300

RESULT 74
Q6P8G4_XENTR
ID Q6P8G4_XENTR PRELIMINARY; PRT; 287 AA.
AC Q6P8G4;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 14.
DE Hypothetical protein MGC75696.
GN Name=MGC75696;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
```

```
RA Butterfield V.S.N., Krywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RA Klein S., Gerhardt D.S.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC EMBL; BC061263; AAH61263.1; -; mRNA.
DR HSSP; P49891; IAOU.
DR Ensemble; ENSXETG0000012821; Xenopus tropicalis.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Hypothetical protein.
SQ SEQUENCE 287 AA; 33145 MW; F5F382031752B4A8 CRC64;

Alignment Scores:
Pred. No.: 1.23e-20 Length: 287
Score: 416.50 Matches: 94
Percent Similarity: 52.2% Conservative: 48
Best Local Similarity: 34.6% Mismatches: 111
Query Match: 9.5% Indels: 19
DB: 4

US-10-768-158-1 (1-2419) x Q6P8G4_XENTR (1-287)
QY 99 CTGGCGCCCTCTGCGCGGGAAGATGAGAGATCGCAACTTCCCGTGGCGCCAGC 158
Db 15 LeuGlyProPheAla--AlaAsnTrpGluAsnValGluLysPheGlnAlaArgProAsp 33
QY 159 GACGTGTGGATCGTCACCTACCTACCCCAAGTCCGGCACCAGCTTGTCTGCAGAGAGTGTCTAC 218
Db 34 AspLeuLeuIleAlaTyrProLysSerGlyThrThrTrpMetSerGluIleValAsp 53
QY 219 TTGGTGACCGAGGGCGTGCACCCGATGAGATCGGCTTGATGAACATCGACGAGCAGCTC 278
Db 54 GlnIleValAlaValSerAsnSerGluArgCysLysThrAlaAlaIleTyrGluArgVal 73
QY 279 CCGGTCTCTGGAGTACCCACACCGC-----GGCTGGACATCATCAAGGAACGTG 326
Db 74 ProPheLeuGluTyrAlaValProAspMetProSerGlyThrGlnAlaLeuAspGlnArg 93
QY 327 ACCTCTCCCGCCCTCATCAAGAGCCACCTGCCCTTTCGCCCTCTCTGACCTCCAC 386
Db 94 AlaSerProArgLeuIleLysThrHisLeuProValGluLeuLeuProLysSerPheTrp 113
QY 387 AATGAGACTCCAAGTTCATCTATATGCTCGCAACCCCAAGGATCTCGTGGTGTCTTAT 446
Db 114 AspAsnLysValLysValIleTyrValAlaArgAsnAlaLysAspValAlaValSerTyr 133
QY 447 TATCAGTTCCACCGCTCTCTCGGACCATGAGTACCGGACCATCTTCAAGAAATTTCTGC 506
Db 134 TyrHisPheTyrArgMetAlaIleValHisProGluProGlyThrTrpAspGluPheLeu 153
QY 507 CGGAGGTTTATGAATGATAGCTGGGCTACCGCTCTCGCTCTGTGGACAGCTGCAGGAGTTC 566
Db 154 AspSerTyrIleAsnGlyLysValCysPheGlySerTrpSerAlaHisValLysGlyTrp 173
QY 567 TGGGAGCACCCGATGAGTTCGAAGCTGCTTTTCTCAAGTATGAAGATCATGATCGGAC 626
Db 174 TrpGlnLysAlaLysGluTrpAspValLeuTyrLeuPheTyrGluAspMetLeuGluAsp 193
QY 627 CTGGTGACAGTGGTGAGCACCTGGCCAGATTCCTGGGGGTCTCTGTGCAAGGCCAG 686
Db 194 ProThrArgGluIleArgLysValLysPheMetGlyLysAspLeuProGluGluThr 213
```

QY 687 CTGGAGCCCTG-----ACGAGACAGTCCACAGCTGGTGGAGCAGCTGCTGCAAC 737  
 Db 214 ValGluLysIleAlaSerGlnThrSerPheLysAlaMetLysGlnAsnGluLeuSerAsn 233  
 QY 738 GCTGAGGCCCTGCCCC-----GTGGCCCGGGGA 764  
 Db 234 TyrSerMetValProSerSerValMetAspHisSerIleSerProPheMetArgLysGly 253  
 QY 765 AGAGTTGGCTGTGGAAGGACATCTCACCGTCTCCATGAATGAGAGTTTGACTGGTG 824  
 Db 254 ValCysGlyAspTrpLysAsnGlnPheThrValAlaGlnAsnGluLysPheAspGluTyr 273  
 QY 825 TATAACAGAGATGGGAAGTGTGACCTCAGCTTT 860  
 Db 274 TyrGlnArgGluMetSerAspGlyAlaLeuSerPhe 285

RESULT 75  
 Q88EV4\_XENLA  
 ID Q88EV4\_XENLA PRELIMINARY; PRT; 276 AA.  
 AC Q68EV4;  
 DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.  
 DT 11-OCT-2004, sequence version 1.  
 DT 07-FEB-2006, entry version 12.  
 DE MGC84291 protein.  
 GN Name=MGC84291;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Brain;  
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.,  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391 (2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Donald M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S.G., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smallos D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Brain;  
 RA Klein S., Gerhard D.S.;  
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NonDerivs License

DR EMBL; BC080096.1; -, mRNA.  
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.  
 DR InterPro; IPR000863; Sulfotransferase.  
 DR Pfam; PF00685; Sulfotransfer\_1; 1.  
 DR ProDom; PD001218; Sulfotransferase; 1.  
 SQ SEQUENCE 276 AA; 31979 MW; 4A2B98654EF33783 CRC64;  
 Alignment Scores:  
 Pred. No.: 1.69e-20 Length: 276  
 Score: 414.50 Matches: 96  
 Percent Similarity: 51.2% Conservative: 48  
 Best Local Similarity: 34.2% Mismatches: 100  
 Query Match: 9.4% Indels: 37  
 Dbs: 2 Gaps: 6  
 US-10-768-158-1 (1-2419) x Q68EV4\_XENLA (1-276)  
 QY 99 CTGCGCCCTTCTGCGCGGGAAGATGGAGAGATGCCAACTTCCCGTGGCGCCAGC 158  
 Db 4 LeuGlyProPheAlaGlu---AsnTrpGluAsnValGluLysPheGlnAlaArgProAsp 22  
 QY 159 GACGTGTGGATCGTCACCTACCCCAAGTCGCGCACCGAGTGTGTCAGGAGTGGTCTAC 218  
 Db 23 AspLeuLeuIleCysThrTyrProLysSerGlyThrThrTrpIleCysGluIleVal--- 41  
 QY 219 TTGTGTGACCGAGCGCTGACCCCGATCGATCGCTTGTGTGTAAC----- 263  
 Db 42 -----AspGlnIleLeuAlaValAsnAsnAlaGluGlyCys 53  
 QY 264 -----ATCGACGACGAGCTCCCGTCTCGGAGTACCCACACGCC----- 302  
 Db 54 LysAsnAlaAlaIlePheGluArgValProPheLeuGluTyrAlaValProAsnMetile 73  
 QY 303 ---GGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCACCTGCCC 359  
 Db 74 SerGlyThrAlaAlaLeuAspGlnArgAlaSerProArgIleIleLysThrHisLeuPro 93  
 QY 360 TACCGCTTTCGCTCTGACCTCCCAATCGAGACTCCAGGTCACATATATGGCTCGC 419  
 Db 94 ValGluLeuLeuProLysSerPheTrpAspAsnLysValLysIleIleTyrValAlaArg 113  
 QY 420 AACCCCAAGGATCTGTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGC 479  
 Db 114 AsnAlaLysAspValAlaValSerTyrThrHisPheTyrGlnMetAlaIleValHisPro 133  
 QY 480 TACCGAGCACCTTCAAGAAATTCGCGGAGGTTTATGAATGATAGTACCTGGGTACCGC 539  
 Db 134 GluProGlyThrTrpAspPheLeuAspSerTyrIleGluGlyLysValCysPheGly 153  
 QY 540 TCCTGGTTTGACGACGTCAGAGTTCCTGGGAGCACCGCATGGAGCTCGAAGCTGCTTTT 599  
 Db 154 ProTrpSerThrHisValLysGlyTyrTrpGlnMetAlaLysLysTrpAspValLeuTyr 173  
 QY 600 CTCAGATGATGAACATGCATCGGACCTGCTGACGATGGTGGAGCAGCTGGCCAGATTC 659  
 Db 174 LeuPheTyrGluAspMetLeuGluAspLeuThrArgGluIleArgLysValLysPhe 193  
 QY 660 CTGGGGGTGTCCTGTGACAAAGCCCGACGTGAA-----GCCCTGACGAGCAGCTGC 710  
 Db 194 MetGlyLysAspLeuSerLysGluLeuValGluLysIleAlaSerLeuThrSerPheLys 213  
 QY 711 CACCGAGCTGGTGACCGAGTGTGCAACGCTGAGCCCTGCGCC----- 752  
 Db 214 AlaMetLysGluAsnLysAsnSerAsnTyrThrThrValProSerSerValMetAspHis 233  
 QY 753 -----GTGGCGCGGGGAAGAGTTGGGTGTGGAGAGACATCTTACCGTC 797  
 Db 234 SerIleSerProPheMetArgLysGlyValCysGlyAspTrpLysAsnGlnPheSerVal 253  
 QY 798 TCCATGAATGAGAGTTTGACTTGTGTATATAACAGAGATGGGAAAGTGTGACCTCAGC 857  
 Db 254 AlaGlnAsnGluLysPheAspLysTyrTyrGlnArgGluMetAlaAspCysAspLeuSer 273

QY 858 TTT 860  
 Db 274 Phe 274

RESULT 76  
 O35401\_MOUSE PRELIMINARY; PRT; 295 AA.  
 AC O35401\_MOUSE  
 DT 01-JAN-1998, integrated into UniProtKB/TrEMBL.  
 DT 01-JAN-1998, sequence version 1.  
 DT 07-FEB-2006, entry version 28.  
 DE Amine N-sulfotransferase.  
 GN Name=Sult1d1; Synonyms=SULT-N;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA MEDLINE=98321187; PubMed=9647753; DOI=10.1006/bbr.1998.8872;  
 RA Sakakibara Y., Yanagisawa K., Takami Y., Nakayama T., Suiko M.,  
 RA Liu M.-C.;  
 RT "Molecular cloning, expression, and functional characterization of  
 RT novel mouse sulfotransferases.";  
 RL Biochem. Biophys. Res. Commun. 247:681-686(1998).  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AF026073; AAC69919.1; -; mRNA.  
 DR PIR; JEO197; JEO197.  
 DR HSSP; P49888; 1HY3.  
 DR Ensembl; ENSMUSG00000029273; Mus musculus.  
 DR MGI; MGI:1926341; Sult1d1.  
 DR GO; GO:0004062; P:aryl sulfotransferase activity; IDA.  
 DR GO; GO:0000103; P:sulfate assimilation; IDA.  
 DR InterPro; IPR000863; Sulfotransferase.  
 DR Pfam; PF00685; Sulfotransfer\_1; 1.  
 DR ProDom; PD001218; Sulfotransferase; 1.  
 KW Transferase.  
 SQ SEQUENCE 295 AA; 35099 MW; 17FE3E02AD6E1269 CRC64;

Alignment Scores:  
 Pred. No.: 2.2e-20 Length: 295  
 Score: 413.00 Matches: 83  
 Percent Similarity: 55.7% Conservative: 63  
 Best Local Similarity: 31.7% Mismatches: 98  
 Query Match: 9.4% Indels: 18  
 DB: 2 Gaps: 2

US-10-768-158-1 (1-2419) x O35401\_MOUSE (1-295)

QY 129 GAGATCGCAACTTCCGGTGGCGCCAGCAGCTGTGGATCGTCACCTACCCCAAGTCC 188  
 Db 30 GlnValGluSerPheGluAlaArgProAspIleLeuIleSerThrTyrrProlysSer 49  
 QY 189 GGCACACGCTGCTCAGAGGTGTCTACTTGGTGAGCCAGCGCGCTGACCCCGATGAG 248  
 Db 50 GlyThrThrTrpValserGluIleLeuAspIleIleTyrAsnAsnGlyAspAlaGluIys 69  
 QY 249 ATCGGCTTTGTAACATCAGCAGCAGCTCCGGTCTCGGAGTACCCACAGCCGGCGCTG 308  
 Db 70 CysLysArgAspAlaileTyrLysArgValProPheMetGluLeuIleIleProGlyIle 89  
 QY 309 -----GACATCATCAGAACTGACCTCTCCCGCTCATCAAGACCCACTG 356  
 Db 90 ThrAsnGlyValGluMetLeuAsnMetProSerProArgIleValIysThrHisLeu 109  
 QY 357 CCCTACCGCTTCTCCCTCTGACCTCCACAATGGAGCTCCAAAGTTCATCTATATGCT 416  
 Db 110 ProValGlnLeuLeuProSerSerPheThrIysAsnAspCysLysIleIleIleTyrValAla 129

QY 417 CGAAACCCCAAGGATCTGGTGTCTTATATATCAGTTCCACCGCTCTCTCGGACCATG 476  
 Db 130 ArgAsnAlaLysAspValValSerTyrTyrPheTyrGlnMetAlaLysIleHis 149  
 QY 477 AGTACCGAGGCACCTTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTGGGTAC 536  
 Db 150 ProGluProGlyThrTrpGluPheLeuGluLysPheMetAlaGlyGlnValSerPhe 169  
 QY 537 GGCTCTCGTTTGGACGACGTCAGAGTTCTGGGAGCACCGCATGGAGCTCGAACTGCTT 596  
 Db 170 GlyProTrpTyrAspHisValLysSerTrpTrpGluLysArgLysGluTyrArgIleLeu 189  
 QY 597 TTTCTCAAGTATGAAGACATCATCGGACCTGTCAGCATGGTGGAGCAGCTGCCACA 656  
 Db 190 TyrLeuPheTyrGluAspMetLysGluAsnProLysCysGluIleGlnGlnIleLeuLys 209  
 QY 657 TTCCTGGGGGTCTCTGTGACAAAGCCAGCTGGAAGCCCTGACGAGCAGCTGCCAC--- 713  
 Db 210 TyrLeuGluLysAspIleProGluGluLeuAsnLysIleLeuTyrHisSerSerPhe 229  
 QY 714 -----CAGCTGGTGGACCACTGTGCTGC 734  
 Db 230 SerValMetLysGluAsnProSerAlaAsnTyrThrThrMetMetLysGluMetAsp 249  
 QY 735 AACCTGAGGCCCTGCGCGCGGGAGAGCTTGGCTGTGGAGGACATCTTCACC 794  
 Db 250 HisSerValSerProPheMetArgLysGlyIleSerGlyAspTrpLysAsnGlnPheThr 269  
 QY 795 GTCTCCATGAATGAGAAGTTTGACTGTGTATAAACAGAGATGGGAAAGTGTGACCTC 854  
 Db 270 ValAlaGlnTyrGluLysPheGluAspTyrValLysLysMetGluAspSerThrLeu 289  
 QY 855 ACCTTT 860  
 Db 290 LysPhe 291

RESULT 77  
 QSEAWO\_XENLA PRELIMINARY; PRT; 303 AA.  
 AC QSEAWO;  
 DT 15-MAR-2005, integrated into UniProtKB/TrEMBL.  
 DT 15-MAR-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 8.  
 DE Hypothetical protein.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; Xenopus; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Egg;  
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Egg;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udelsin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.",  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Egg;  
 RC Klein S., Gerhard D.S.;  
 RA Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; BC090224; AAH90224.1; -; mRNA.  
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.  
 DR InterPro; IPR000863; Sulfotransferase.  
 DR Pfam; PF00685; Sulfotransferase\_1; 1.  
 DR ProDom; PD001218; Sulfotransferase; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 303 AA; 35647 MW; B940559678751DA1 CRC64;  
 Alignment Scores:  
 Pred. No.: 2,21e-20 Length: 303  
 Score: 413.00 Matches: 89  
 Percent Similarity: 52.0% Conservative: 55  
 Best Local Similarity: 32.1% Mismatches: 115  
 Query Match: 9.4% Indels: 18  
 DB: 2 Gaps: 3  
 US-10-768-158-1 (1-2419) x QSEAWO\_XENLA (1-303)  
 QY 84 TTCATGCGCGCTCGGCTCGCGCTTCTGCGCGGGAAGATGAGGAGATCGCCAACTTC 143  
 DB 23 PheArgGlyValProMetAlaLysMetAsnValGluAsnTrpGluProValGluTrpPhe 42  
 QY 144 CCGGTGCGCGCCAGCGAGCTGTGGATCTGCTACCTACCCCAAGTCCGCGCAGCTGTGCT 203  
 DB 43 GlnAlaArgHisAspValValLeuAlaThrTyrProLysAlaGlyThrTrpVal 62  
 QY 204 CAGAGGTGTCTACTTGTGTAGCCAGCGCGCTGACCCCGATGAGATCGGCTGTATGAC 263  
 DB 63 SerGluLeuLeuAspMetIleTyrAsnGlyGlyAspLeuGluLysCysGlnArgAspThr 82  
 QY 264 ATCGACGACGAGCTCCCGTCTCTGAGTACCCACAGCGCGGCTG-----GAC 311  
 DB 83 IleTyrAsnArgValProTyrMetGluLeuArgIleProGlyMetProSerGlyValAsp 102  
 QY 312 ATCATCAAGGAACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTCTG 371  
 DB 103 GlnLeuGluLeuLeuAlaSerProArgLeuIleLysThrHisLeuProIleGlnLeuMet 122  
 QY 372 CCCTCTGACCTCCCAATGGAGATCCCAAGTCTATATGCTGCGCAACCCCAAGAT 431  
 DB 123 ProGluSerPheTrpGluLysLysCysValIleTyrValAlaArgAsnAlaLysAsp 142  
 QY 432 CTGGTGGTGTCTATTATACAGTTCCACCGCTCTCTGCGGACCATGAGTACCGAGGCACC 491  
 DB 143 ValAlaValSerTyrPhePheHisGlnMetValLysAlaLeuProAspProGlyPro 162  
 QY 492 TTTCAGGAATCTTCGCGGAGTTTATGAATGATGAAGCTAGGCTACCGCTCTCTGTTGAG 551  
 DB 163 TrpAspLysPheLeuAlaAspTyrMetAsnGlyThrValSerTyrGlySerTrpTyrAsp 182  
 QY 552 CACGTGCAGGAGTCTTCGGAGCACCGCATGGACTCGAAGCGCTCTTTCTCAAGTATGAA 611  
 DB 183 HisValLysGlyTyrTrpGluLysArgLysGlnTyrGlyIleLeuTyrLysPheTyrGlu 202  
 QY 612 GACATGCATCGGACCTGTGTACGATGGTGGAGCAGCTGGCCAGATTCCTCGGGGTGTC 671

DB 203 AspLeuLysGluAspProLysArgGluLysLysLysLeuHisTyrLeuGluLysGlu 222  
 QY 672 TGTGACAAAGCCAGCTGGAGACCTGACGAGCAGTCCACACAGCTGGTG----- 722  
 DB 223 LeuSerAspGluValLeuGluLysIleValHisThrSerPheGlnValMetSerLys 242  
 QY 723 GACACAGTGTCTCAACGCTGAGGCGCTGCGC----- 752  
 DB 243 AsnAspMetAlaAsnTyrLysThrIleProAsnAspIleLeuAsnGlnThrAsnThrAla 262  
 QY 753 ---GTGGCGCGGGAAGATGTGGCTGTGGAGACATCTTCACGCTCTCCATGAATGAG 809  
 DB 263 PheMetArgLysGlyGluAlaGlyAspTrpLysAsnHisPheThrValAlaGlnAsnGlu 282  
 QY 810 AAGTTTCACTGGTGTATTAACAGAGATGGGAAGATGTGACCTCAGCTTT 860  
 DB 283 AlaPheAspAsnGlnTyrGlnGluMetLeuGlyThrSerLeuHisPhe 299  
 RESULT 78  
 Q80VR3 MOUSE  
 ID Q80VR3 MOUSE PRELIMINARY; PRT; 304 AA.  
 AC Q80VR3;  
 DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.  
 DT 07-FEB-2006, entry version 18.  
 DE Sulfotransferase family, cytosolic, 1C, member 1.  
 GN Name=Sult1c1;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Olfactory epithelium;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Berge J.G.,  
 RA Klausner R.F., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.D., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.",  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Olfactory epithelium;  
 RC Director MGC Project;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; BC045149; AAH45149.1; -; mRNA.  
 DR HSP; P50224; 1CJM.  
 DR Ensembl; ENSMUSG0000023943; Mus musculus.  
 DR MGI; MGI:102928; Sult1c1.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0004062; P:aryl sulfotransferase activity; IEA.  
 DR GO; GO:0006790; P:sulfur metabolism; IEA.

DR	InterPro: IP0000863; Sulfotransferase.	DR	Pfam: PF00685; Sulfotransferase_1.	DR	ProDom: PD001218; Sulfotransferase; 1.	KW	Transferase.	Q	SEQUENCE 304 AA; 35797 MW; 9CF4A4B63713B977 CRC64;
Alignment Scores:									
Pred. No.:	2,4e-20	Length:	304						
Score:	412.50	Matches:	89						
Percent Similarity:	52.0%	Conservative:	56						
Best Local Similarity:	31.9%	Mismatches:	115						
Query Match:	9.4%	Indels:	19						
DB:	2	Gaps:	3						
US-10-768-158-1 (1-2419) x Q80VR3_MOUSE (1-304)									
Qy	81	GAGTTCATGCGTGC	GCGCTCGCCGCTCTCTCCCGCGGAAGATGAGGAGATGCCCAAC	140					
Db	22	GluValAsnGlyIleLeuMetSerLysMetMetSerGluAsnTrpAspLysIleTrpAsn	411						
Qy	141	TTCCCGGTGCGCCGACGAGCGTGTGGATCGTCACTACCCACCAAGTCGGGACCACTGTG	200						
Db	42	PheGlnAlaLysProAspAspLeuLeuIleLaThrTyRAlaLysAlaGlyThrThrTrp	61						
Qy	201	CTGACGAGGTGCGTCTACTGTGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATG	260						
Db	62	ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla	81						
Qy	261	AACATCGACGAGCAGCTCCCGTCTCGGAGTACCCACACAGCCG	305						
Db	82	AsnThrTyRAspArgHisProPheIleGluThrLeuProProLeuAsnSerGly	101						
Qy	306	CTGCACATCATCAAGGAACCTGACCTCTCCCGCCTCATCAGAGCCACCTGCGCTACCGC	365						
Db	102	LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValGln	121						
Qy	366	TTTCTGCCCTCGACCTCCCAATGAGACTCCAAGGTCACTATATGCTCGCAACCCC	425						
Db	122	MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyRValAlaArgAsnAla	141						
Qy	426	AAGATCTGGTGGTGTCTATTATCATGTTCCACCGCTCTCTCCGACCATGAGCTACCGA	485						
Db	142	LysAspCysLeuValSerTyRTrpPheSerArgMetAsnLysMetLeuProAspPro	161						
Qy	486	GGCACCTTTCAAGAAATCTCGCGGAGGTTTATGAATGATGAAGCTGGGCTACGGCTCTCGG	545						
Db	162	GlyThrLeuGlyGluTyRileGluThrPheLysAlaGlyLysValLeuTrpGlySerTrp	181						
Qy	546	TTTGAGCAGTGCAGGAGTCTCGGAGCACCGCATGGATCGAAGCTGTTTTTCTCAAG	605						
Db	182	TyrAspHisValLysGlyTrpTrpAspValLysAspLysHisArgIleLeuTyRLeuPhe	201						
Qy	606	TATCAAGACATGCATCGGGACCTGTGTGACGATGGTGGAGCAGCTGCCAGATTTCTGGGG	665						
Db	202	TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleValLysPheLeuGlu	221						
Qy	666	GTGTCTGTGCAAGAGCCACCTGGAAGCCCTGACGAGCAGCTGCCAGCTGGTG---	722						
Db	222	LysAspIleSerGluGluValLeuAsnLysIleIleHisHisThrSerPheAspValMet	241						
Qy	723	-----GNCCAGTGTGCAACGCTGAGCGCTGCCCC-----	752						
Db	242	LysGlnAsnProMetAlaAsnTyRThrThrLeuProSerSerIleMetAspHisSerIle	261						
Qy	753	-----GTGGGCGGGGAAGAGTGTGGGTGTGGAAGGACATCTTCACTGCTCCCATG	803						
Db	262	SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyRPhThrValAlaGln	281						
Qy	804	AATGAGAAGTTTGACTTGGTGTATATAACAGAAAGATGGGAAAGTGTGACCTACGTTT	860						
Db	282	SerGluAspPheAspGluAspTyRArgLysLysMetAlaGlySerThrIleThrPhe	300						

RESULT 79

```

US-10-768-158-1 (1-2419) x ST1A1_BOVIN (1-294)
QY 69 AGCAAGTACTTCGATTCATCGCGTGGCGTGGCGCCCTTCTCCCGCGGAAGATGGAG 128
   ::::::::::::::::::::
Db 12 AlalysytrVallys-----GlyleProleuilellystrPheAlaGluAlaLeuGly 29
   ::::::::::::::::::::
QY 129 GAGATCGCAACTCCCGGTGGCGGCCAGCAGCGTGTGGTGTGGAGCGCGCTGACCCCGATGAG 188
   ::::::::::::::::::::
Db 30 ProLeuGluSerPheGluAlaTrpProAspAspLeuilellySerThrTrpProLysSer 49
   ::::::::::::::::::::
QY 189 GGCACACAGCTGCTGCGAGAGTGTCTACTTGTGTGGAGCGCGCTGACCCCGATGAG 248
   ::::::::::::::::::::
Db 50 GlyThrThrTrpValSerGluileLeuAspLeuilellyTrpGlnGluGlyAspLeuGluLys 69
   ::::::::::::::::::::
QY 249 ATCGGCTGTGTAACATCGACGAGCGTCCCGGTCCTGGAGTACCCACACGCG----- 302
   ::::::::::::::::::::
Db 70 CysGlnArgAlaProValPheLeuArgValProPheLeuGluPheSerAlaProGlyVal 89
   ::::::::::::::::::::
QY 303 -----GGCTGGACATCATCAAGAACTGACCTCTCCCGCCTCATCAAGAGCCACTG 356
   ::::::::::::::::::::
Db 90 ProThrGlyValGluLeuLeuLysAspThrProAlaProArgLeuLeuLysThrHisLeu 109
   ::::::::::::::::::::
QY 357 CCCTACCGCTTCTGCTCTGCTCCACATCGAGACTCCAGGTCTATATATGCT 416
   ::::::::::::::::::::
Db 110 ProLeuAlaLeuProLysThrLeuLeuAspGlnLysValLysVallellyleAla 129
   ::::::::::::::::::::
QY 417 CGCAACCCCAAGGATCTGGTGTGTTATTATCATGTTCCACCGCTCTCTCGGACCATG 476
   ::::::::::::::::::::
Db 130 ArgAenAlaLysAepValAlaValSerTyrrHisPheTyArgMetAlaLysValHis 149
   ::::::::::::::::::::
QY 477 AGCTACCGAGCACCTTTCAAGATTCTCGCGAGGTTTATGATGATGAAGTGGCTAC 536
   ::::::::::::::::::::
Db 150 ProAspProGlyThrTrpAspSerPheLeuGluLysPheMetAlaGlyGluValCysTr 169
   ::::::::::::::::::::
QY 537 GGCTCTCTGTTGACACGCTGAGGAGTCTGGGAGACCCGATGCGACTCGAACTGCTT 596
   ::::::::::::::::::::
Db 170 GlySerTrpTyrrGlnHisValGlnGluTrpGluLeuSerHisThrHisProValLeu 189
   ::::::::::::::::::::
QY 597 TTTCTCAAGTATGAGACATCATCGGACCTGCTGGTGTGACGATGGTGGAGCAGCTGCCAGA 656
   ::::::::::::::::::::
Db 190 TyrLeuPheTyrrGluAspIleLysGluAspProLysArgGluileGlnLysIleLeuGlu 209
   ::::::::::::::::::::
QY 657 TTCTCTGGGGTGTCTGTGACAGCCGCTGGAAGCCCTG-----ACGAGCAC 707
   ::::::::::::::::::::
Db 210 PheIleGlyArgSerLeuProGluGluThrValAspHisIleValGlnArgThrSerPhe 229
   ::::::::::::::::::::
QY 708 TGCCACCAAGCTGTGTGACACGCTGTCACACGCTGAGCGCCCTGCGCGTGGC----- 758
   ::::::::::::::::::::
Db 230 LysGluMetLysLysAenProMetThrAsnTyrrSerThrIleProThrAlaValMetAsp 249
   ::::::::::::::::::::
QY 759 -----CGGGAGAGATTTGGCTGTGGAGAGCATCTTACCC 794
   ::::::::::::::::::::
Db 250 HisSerIleSerAlaPheMetArgLysGlyleThrGlyAspTrpLysSerThrPheThr 269
   ::::::::::::::::::::
QY 795 GTCTCCCATGATGAGAACTTTGACTTGTGTATAAACAGAGATG 839
   ::::::::::::::::::::
Db 270 ValAlaGlnAenGluLeuPheGluAlaHisTyrrAlaLysLysMet 284
   ::::::::::::::::::::

RESULT 80
Q7S293_XENLA PRELIMINARY; PRT; 294 AA.
ID Q7S293_XENLA
AC 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DR 07-FEB-2006, entry version 14.
DE MGC64389 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN (1)

```

## NUCLEOTIDE SEQUENCE.

```

RP TISSUE=Whole;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Halle S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]

```

## NUCLEOTIDE SEQUENCE.

```

RP TISSUE=Whole;
RC MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391(2002).
[3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX Klein S., Strausberg R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; BC053792; AAH53792.1; -; mRNA.
DR HSP; P49888; 1HV3.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
SQ SEQUENCE 294 AA; 34217 MW; 9DA265CAE40BE476 CRC64;

```

## Alignment Scores:

Pred. No.:	2.58e-20	Length:	294
Score:	412.00	Matches:	91
Percent Similarity:	51.0%	Conservative:	58
Best Local Similarity:	31.2%	Mismatches:	121
Query Match:	9.4%	Indels:	22
DB:	2	Gaps:	4

US-10-768-158-1 (1-2419) x Q7S293\_XENLA (1-294)

```

QY 57 GGGGAGTTCGAG-----AGCAAGTACTTCAGTTCCATGTCGTCGCGTCCGCC 107
   |||||
Db 3 GlyGluLeuGluTrpArgLysAspTrpValAspValTyrrGlyIleProMetIleAla 22
   |||||
QY 108 TTCTGCCCGGGAAGATGGAGAGATCGCAACTTCCCGTGGCGGCCAGGAGTGTGG 167
   |||||
Db 23 AlaPheSerSerAenTrpGluArgIleLysAenPheGlnAlaArgAlaAspIleVal 42
   |||||
QY 168 ATCTGTACCTACCCCAAGTCGCGACCACTGTCGTCGAGGAGTGGTCTACTTGTGAGC 227
   |||||
Db 43 IleCysThrTyrrProLysSerGlyThrTrpIleSerGluIleValAspValIleVal 62
   |||||
QY 228 CAGGCGCTGACCCCGATGAGATCGGCTTGTATGAACATGACGAGAGCTCCCGTCTG 287
   |||||

```









```
RA Director MGC Project;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC066190; AAH66190.1; -; mRNA.
DR HSP; P49891; LAOU.
DR Ensembl; ENSMUSG0000029273; Mus musculus.
DR MGI; MGI:1926341; Sult1d1.
DR GO; GO:0004062; F:aryl sulfotransferase activity; IDA.
DR GO; GO:000103; P:sulfate assimilation; IDA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 295 AA; 35011 MW; A569PAP60CC0716F CRC64;

Alignment Scores:
Pred. No.: 2 59e-20 Length: 295
Score: 412.00 Matches: 83
Percent Similarity: 55.7% Conservative: 63
Best Local Similarity: 31.7% Mismatches: 98
Query Match: 9.4% Indels: 18
DB: 2 Gaps: 2

US-10-768-158-1 (1-2419) x QN2D1_MOUSE (1-295)
QY 129 GAGATCGGCAATCTCCCGTGGCGCCAGCAGTGTGTGATCGTCACCTACCCAGTCC 188
DB 30 GlnValGluSerPheGluAlaArgProAspAlaLeuLeuSerThrTyProLysSer 49
QY 189 GGCACACAGTCTGCTGAGAGTGTCTACTTGTGAGCCAGCGGCGTGCACCGATAG 248
DB 50 GlyThrTrpValSerGluLeuLeuAspLeuLeuTyAsnAsnGlyAspAlaGluLys 69
QY 249 ATCGGCTTGATGAACATGACGACGAGCTCCGGTCTTGAGTACCCAGCGGCGCTG 308
DB 70 CysLysArgAspAlaLeuTyLysArgValProPheMetGluLeuLeuLeuProGlyLe 89
QY 309 -----GACATCATCAAGAACTGACTCTCCCGCTCTCAAGAGCCACCTG 356
DB 90 ThrAsnGlyValGluMetLeuAsnMetProSerProArgIleValLysThrHisLeu 109
QY 357 CCCTACCGCTTCTCCCTCTGACCTCCACATGAGACTCCAGTCTATATAGCT 416
DB 110 ProValGlnLeuLeuProSerPheThrLysAsnAspCysLysLeuLeuLeuValAla 129
QY 417 CGCAACCCCAAGATCTGTGTGTCTTATTATAGTCCACCGCTCTCTGGGACCATG 476
DB 130 ArgAsnAlaLysAspValValSerTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy 149
QY 477 AGCTACCGAGGACCTTTCAAGAATTCTCCCGAGGTTTATGAATGATAAGTGGCTAC 536
DB 150 ProGluProGlyThrTrpGluGluPheLeuGluLysPheMetAlaGlyGlnValSerPhe 169
QY 537 GCCTCCTGTTGACGACGTGAGAGTCTGGAGACCGGATCGACTCGACCTGCTT 596
DB 170 GlyProTrpTyAspHisValLysSerTrpTrpGluLysArgLysGluTyArgileLeu 189
QY 597 TTCTCAAGTATGAAGACATGATCGGACCTGGTGCAGTGTGTGAGAGTGGCCAGA 656
DB 190 TyrLeuPheTyGluAspMetLysGluAsnProLysCysGluLeuGlnLysLeuLys 209
QY 657 TTCTCGGGGTCTCTGTGACAGGCCAGCTGGAGCCCTGACGAGCACTGCCAC--- 713
DB 210 PheLeuGluLysAspIleProGluLeuLeuAsnLysLeuLeuTyHisSerPhe 229
QY 714 -----CAGCTGGTGGACAGTCTGCT 734
DB 230 SerValMetLysGlyAsnProSerAlaAsnTyThrThrMetMetLysGluGluMetAsp 249
QY 735 AACGCTGAGGCCCTGCCCGTGGCGGAGAGTGTGGCTGTGGAAGGACATCTTCACC 794

Db 250 HisSerValSerProPheMetArgLysGlyLeuSerGlyAspTrpLysAsnGlnPheThr 269
QY 795 GTCTCCATGAATGAGAGTTTGACTTGTGTATATAACAGAGATGGGAAAGTGTGACCTC 854
DB 270 ValAlaGlnTyGluLysPheGluGluAspTyValLysLysMetGluAspSerThrLeu 289
QY 855 ACGTTT 860
DB 290 LysPhe 291

RESULT 84
Q921G0_RAT PRELIMINARY; PRT; 308 AA.
AC Q921G0;
DT 01-MAY-1999, integrated into UniProtKB/TrEMBL.
DT 01-MAY-1999, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Tyrosine-ester sulfotransferase (EC 2.8.2.9).
GN Name=Sult1d1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
RL Herzmann A., Stoffel W.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U32372; AAC99890.1; -; mRNA.
DR HSP; P49888; 1HY3.
DR Ensembl; ENSRNOG000001960; Rattus norvegicus.
DR RGD; 620491; Sult1d1.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0017067; F:tyrosine-ester sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 308 AA; 36341 MW; 68719C9223BE9EAE CRC64;

Alignment Scores:
Pred. No.: 2 62e-20 Length: 308
Score: 412.00 Matches: 93
Percent Similarity: 52.8% Conservative: 59
Best Local Similarity: 32.3% Mismatches: 110
Query Match: 9.4% Indels: 26
DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x Q921G0_RAT (1-308)
QY 63 TTCGAGACCAAGTACTTCGAGTTCGAGTTCGCGGTGCGGCTCTTCGCGCGGGAAG 122
DB 8 PheArgArgGluLeuValAspValGlnGlyLeuProLeuPheTrpSerIleAlaGluLys 27
QY 123 ATGGAGGAGATCGCAACTTCGCGGTGCGGCGCCAGCAGCTGTGGATCGTCACCTACCCC 182
DB 28 TrpSerGlnValGluSerPheGluAlaArgProAspAlaLeuLeuSerThrTyPro 47
QY 183 AAGTCCGCAACAGTGTCTGACGAGTGTCTACTTGTGTGAGCCAGGCGGTGACCCC 242
DB 48 LysSerGlyThrThrTrpIleSerGluLeuLeuAspLeuLeuTyAsnAsnGlyAspAla 67
QY 243 GATGAGATCGGCTTGATGAACATGACGAGGAGCTCCCGGTCCCTGAGTACCCAGCCG 302
DB 68 GluLysCysLysArgAspAlaIleTyArgArgValProPheMetGluLeuLeuLeu 87
QY 303 GCGCTG-----GACATCATCAAGGAAGTACCTCTCCCGCTCATCAAGAGC 350
```

```

Db 88 GlyIleThrAsnGlyValGluMetLeuAspAsnMetGlnSerProArgLeuValIleThr 107
QY 351 CACCTGCCCTACCGCTTTCTGCCCTCTGACCTCCACAAATGGAGACTCCAAAGGTCTATCTAT 410
Db 108 HisLeuProValGlnLeuLeuProSerSerPheTyrArgAsnAspCysLysMetIleTyr 127
QY 411 ATGGCTCCGAACCCCAAGATCTGGTGGTCTTATTATCATGTTCCACCGCTCTCTGGGG 470
Db 128 ValAlaIleArgAsnAlaLysAspValAlaValSerTyrTyrPheHisGlnMetAlaLys 147
QY 471 ACCATGAGCTACCGAGGACCTTTCAAGAATCTCCGGAGGTTTGAATGAATGAAGCTG 530
Db 148 MetHisProGluProGlyThrTrpGluGluPheLeuGluLysPheMetAlaGlyGlnVal 167
QY 531 GGCTACGGCTCTCTGTTTGAACAGCTGAGGAGTTCTGGGACACCGCATGGAATCGAAC 590
Db 168 SerPheGlyProTyrAspHisValLysGlyTyrTrpGluLysArgLysGluTyrArg 187
QY 591 GTGCTTTTCTCAAGTATGAACATATGATCGGACCTGGTGACATGGTGGAGCAGCTG 650
Db 188 IleLeuTyrCysPheTyrGluAspMetLysGluAspProLysCysGluIleGlnLysVal 207
QY 651 GCCAGATTCTG----- 662
Db 208 LeuLysPheLeuGluLysAspIleProGluGluValAlaAsnLysIleLeuTyrHisSer 227
QY 663 GGGGTGTCTGTGACAAAGCCCGAGCTGGAAGCC---CTGACGAGCAGCTGCCACAGCTG 719
Db 228 SerPheSerValMetLysAlaAsnProSerAlaAsnTyrThrMetMetLysGluGlu 247
QY 720 GTGGACCAAGTGTGCAAGCTGAGCCCTGCCCC---GTGGCCGGGGAGAGTTGGGCTG 776
Db 248 MetAspGlnSerValSer-----ProPheMetArgLysGlyIleSerGlyAsp 263
QY 777 TGGAGGACATCTTCACCGCTCCATGAATGAGAGTTTGACTTGGTGTATTAACAGAG 836
Db 264 TrpLysAsnGlnPheThrValAlaGlnTyrGluLysPheGluLysAspTyrValLysLys 283
QY 837 ATGGGAAAGTGTGACCTCACGTTT 860
Db 284 MetGluGluSerThrLeuLysPhe 291

RESULT 85
ID Q9R2C2 MOUSE PRELIMINARY; PRT; 309 AA.
AC Q9R2C2
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DE 07-FEB-2006, entry version 21.
DE Tyrosine-ester sulfotransferase (EC 2.8.2.9).
GN Name=Sult1d1;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BALB/c; TISSUE=Kidney;
RA Herrmann A., Stoffel W.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DB EMBL; U32371; AAC99889.1; -; mRNA.
DR HSSP; P49888; 1HY3.
DR Ensembl; ENSMUSG00000029273; Mus musculus.
DR MGI; MGI:1926341; Sult1d1.
DR GO; GO:0004062; F:aryl sulfotransferase activity; IDA.
DR GO; GO:0000103; P:sulfate assimilation; IDA.
DR InterPro; IPR0000863; Sulfotransferase.

```

```

DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ SEQUENCE 309 AA; 36713 MW;  E5D5E1CD3C6D5117 CRC64;

Alignment Scores:
Pred. No.: 2,62e-20 Length: 309
Score: 412.00 Matches: 83
Percent Similarity: 55.7% Conservative: 63
Best Local Similarity: 31.7% Mismatches: 98
Query Match: 9.4% Indels: 18
DB: 2 Gaps: 2

US-10-768-158-1 (1-2419) x Q9R2C2_MOUSE (1-309)

QY 129 GAGATCGCAACTTCCCGTGGCGCCAGCCAGCTGTGGATCGTCACCTACCCCAAGTCC 188
Db 30 GlnValGluSerPheGluAlaArgProAspAspIleLeuIleSerThrTyrProLysSer 49
QY 189 GGCACCACTGTGTGAGGAGTGTCTACTTGTGTGAGCCAGGGCGCTGACCCCATGAG 248
Db 50 GlyThrThrTrpValSerGluIleLeuAspLeuIleTyrAsnAsnGlyAspAlaGluLys 69
QY 249 ATCGGCTTGATGAACATCGACGAGCAGCTCCGGTCTCGGAGTACCCACAGCCGGGCTG 308
Db 70 CysLysArgAspAlaIleTyrLysArgValProPheMetGluLeuIleIleProGlyIle 89
QY 309 -----GACATCATCAAGGAACCTGACCTCTCCCGGCTCATCAAGAGCCACCTG 356
Db 90 ThrAsnGlyValGluMetLeuAsnAsnMetProSerProArgIleValLysThrHisLeu 109
QY 357 CCTTACCGCTTTCTGCTCTGAGCTCCACAAATGGAGACTCCCAAGTCACTATATGGCT 416
Db 110 ProValGlnLeuLeuProSerSerPheTrpLysAsnAspCysLysIleIleTyrValAla 129
QY 417 CGCAACCCCAAGGATCTGGTGTCTTATTTATCATGTTCCACCGCTCTCTCGGACCATG 476
Db 130 ArgAsnAlaLysAspValValSerTyrTyrPheTyrGlnMetAlaLysIleHis 149
QY 477 AGCTACCGAGGACCTTTCAAGAATTTCTGCGGAGGTTTATGAATGATAAGCTGGGCTAC 536
Db 150 ProGluProGlyThrTrpGluGluPheLeuLysPheMetAlaGlyGlnValSerPhe 169
QY 537 GGCTCTGGTTTGTGACGACGTGACGAGTTCTGGGAGCACCGCATGGACTCGAACGTGTT 596
Db 170 GlyProTyrAspHisValLysSerTrpTrpGluLysArgLysGluTyrArgIleLeu 189
QY 597 TTTCTCAAGTATGAACATCATCGGACCTGGTGTGACGATGGTGGAGCAGCTGCCCCAGA 656
Db 190 TyrLeuPheTyrGluAspMetLysGluAsnProLysCysGluIleGlnLysIleLys 209
QY 657 TTCTCTGGGGTGTCTGTGACAAAGCCCGCTGGAAGCCCTGACGGAGCAGCTGCCAC--- 713
Db 210 PheLeuGluLysAspIleProGluGluLeuAsnLysIleLeuTyrHisSerPhe 229
QY 714 -----CAGCTGTGGACACAGTGTGC 734
Db 230 SerValMetLysGluAsnProSerAlaAsnTyrThrMetLysGluGluMetAsp 249
QY 735 AACGCTAGGCCCTGCGCGTGGCGGGAAGAGTTGGCTGTGGAGGACACTCTTACC 794
Db 250 HisSerValSerProPheMetArgLysGlyIleSerGlyAspTrpLysAsnGlnPheThr 269
QY 795 GTCTCCCAATGAAGATTTGACTTGGTGTATTAACAGAGATGGGAAGTGTGCACCTC 854
Db 270 ValAlaGlnTyrGluLysPheGluLysAspTyrValLysLysMetGluAspSerThrLeu 289
QY 855 ACCTTT 860
Db 290 LysPhe 291

RESULT 86
Q66KW4_XENLA

```

Db           9 LeuAlaGluMetGluAsnPheGlnValThrMetGlyHisIleGlu----- 24  
 Qy         81 GAGTTCCATGGCGTGC GGCTGCCG---CCCTTCGTCCCGGGAAGATGAGAGAGATCGCC 137  
             |||||           |||||           |||           :::   ::: ||||  
 Db         25 -----GlyValProLeuProGlyProThrCys---AspGluTrpAspThrIleTyR 40  
 Qy         138 AACTTCCGGTGCGGCCAGGACGAGTGTGGATCGTCACTACCACCAAGTCCGGCACACG 197  
             |||||           |||           |||           |||||           |||:  
 Db         41 AsnPheGlnAlaArgGluAspApileLeuIleAlaThrTyProLysSerGlyThrThr 60  
 Qy         198 TTGCTGAGGAGGTGGTCTACTTGGTAGAGCAGGGCGCTGACCCCGATCGATCGCTTG 257  
             ::: |||           |||           |||           |||           |||:  
 Db         61 TrpMetGlnGluIleValAspLeuIleLeuGlnGluGlyAspValGlnLysSerMetArg 80  
 Qy         258 ATGAACATCGACGACGAGCTCCCCTCTCTGAGTAGTACCCACACAGCCG----- 302  
             ::: |||           |||           |||           |||           |||:  
 Db         81 AlaProCysPheIleLysValProPheIleGluMetIleProProLysSerMetProSer 100  
 Qy         303 GGCCTGACATCATCAAGAAACTGACCTCTCCCGCCTCATCAAGAGCACCTGCGCTTAC 362  
             |||||           |||           |||           |||           |||:  
 Db         101 GlyLeuGluLeuAlaLysThrMetLysSerProArgIleLeuLysThrHisLeuProIle 120  
 Qy         363 CGCTTTCGCCCTCTGACCTCCACAATGGAGACTCCAGAGTCATCTATATGCTCGCAAC 422  
             |||||           |||           |||           |||           |||:  
 Db         121 AsnLeuLeuProProSerPheTrpGluLysAsnAlaLysValValTyValAlaArgAsn 140  
 Qy         423 CCCAAGATCTGGTGGTGTCTATTATTCAGTTCCACCGCTCTCTCGGACCATGAGTAC 482  
             |||||           |||           |||           |||           |||:  
 Db         141 AlalysaspCysMetValSerTyTyTrpPheHisLysMetAsnThrPheLeuLeuAsp 160  
 Qy         483 CGAGGCACCTTCAAGAATCTCGCGGAGTTTATGAATATATACTGGGTACCGCTCC 542  
             |||||           |||           |||           |||           |||:  
 Db         161 ProGlyThrTrpAspAsnPhePheSerGluPheLeuSerGlyAspValProTrpGlySer 180  
 Qy         543 TGGTTTGAGCACCTGACGAGGTTCTGGAGACACCGCATGGACTCGAACGTGCTTTTCTC 602  
             |||||           |||           |||           |||           |||:  
 Db         181 TrpPheAspHisValLeuGlyTrpTrpLysAlaMetAspLysHisGlnIleLeuPheIle 200  
 Qy         603 AAGTATGAAGACATGATCGGGACCTGGTGACGATGGTGGAGCACCTGCCAGATTCCTG 662  
             |||||           |||           |||           |||           |||:  
 Db         201 PheTyrgluAspMetIleGluAspPrometArgGluIleArgLysValMetLysPheLeu 220  
 Qy         663 GGGGTGTCTGTGACAAGGCCACGCTGGAAAGCCCTGACGGAGCAGTGGCCAC----- 713  
             |||||           |||           |||           |||           |||:  
 Db         221 GlyLysAspLeuSerAspGluAlaLeuGluAsnVallylTyHisSerPheGlnAla 240  
 Qy         714 CAGCTGTGGACCAAGTGTGTCGACCGCTGAGCCCTGCCCC----- 752  
             |||           |||           |||           |||           |||:  
 Db         241 MetLysGluAsnPrometThrAsnAsnSerThrValProAsnSerIleMetAspAspThr 260  
 Qy         753 -----GTGGGCGGGGAAGAGTGTGGCTGTGGAGGACATCTTACCCTGCTCC 800  
             |||           |||           |||           |||           |||:  
 Db         261 IleSerPropheMetArgLysGlyIleValGlyAspTrpLysThrHisPheSerValThr 280  
 Qy         801 ATGAATGAGAAGTTTGACTTGGTGTATAAACAAGATGGGAAGAGTGTGACCTACGTTT 860  
             |||           |||           |||           |||           |||:  
 Db         281 GlnAsnPheIlePheAspLysGluTyrylLysLysMetGluGlySerGlyLeuAsnPhe 300

RESULT 87  
 Q5HZUO\_XENTR PRELIMINARY; PRT; 297 AA.  
 AC Q5HZUO;  
 DT 15-FEB-2005, integrated into UniProtKB/TrEMBL.  
 DT 15-FEB-2005, sequence version 1.  
 DE 07-FEB-2006, entry version 10.  
 DE Hypothetical LOC496998.  
 GN Name=LOC496998;  
 OS Xenopus tropicalis (Western clawed frog) (*Xenopus tropicalis*).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 CC Xenopodinae; Xenopus; *Silurana*.  
 NCBI\_TaxID=8364;  
 RN [1]

RP NUCLEOTIDE SEQUENCE.  
 RX TISSUE=whole body;  
 RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny K.B., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=whole body;  
 RA Klein S., Gethard D.S.;  
 RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.  
 CC  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NonDerivs License  
 CC  
 DR EMBL; BC088888; AAH88888.1; -; mRNA.  
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.  
 DR InterPro; IPR000863; Sulfotransferase.  
 DR Pfam; PF00685; Sulfotransferase\_1; 1  
 DR ProDom; PD001218; Sulfotransferase; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 297 AA; 34863 MW; E8248162DC31D7D2 CRC64;  
 Alignment Scores:  
 Pred No.: 3,31e-20 Length: 297  
 Score: 410.50 Matches: 85  
 Percent Similarity: 55.7% Conservative: 61  
 Best Local Similarity: 32.4% Mismatches: 99  
 Query Match: 9.3% Indels: 17  
 DB: 2 Gaps: 4  
 US-10-768-158-1 (1-2419) x Q5HZU0\_XENTR (1-297)  
 Qy 120 AAGTGGAGGAGATGCCAACTTCCTCCGGTGGCGCCAGCAGCTGGTGGATCGTCACTAC 179  
 Db 30 LysileAspSerIleGlnAspPheLysValLysAspThrAspValPheLeuValThrTyr 49  
 Qy 180 CCCAAGTCCGACGAGCTTCTGTCAGGAGGTGGTCTACTTGGTG---AGCCAGGCGCT 236  
 Db 50 ProlysThrGlyThrIleTrpThrGlnGlnLeuSerLeuIlePheAsnGluGlyHis 69  
 Qy 237 GACCCGATGATCGGCTTGATGACATCGACGAGCTCCCGTCTCGAGTACCCA 296  
 Db 70 ArgAsnGlyThrGluAlaIleAlaAsnVal---PheArgValProTrpIleGluTyrThr 88  
 Qy 297 CAGCGCGCCTGGACATCATCAAGAACTGACCTCTCCCGCCTCATCAAGACCCACCTG 356  
 Db 89 HisSerLysValAsp---TyrAspSerArgProSerProArgLeuPheSerSerHisLeu 107  
 Qy 357 CCTTACCGCTTTCGCTCTGACCTCCACATGAGAGCTCCAGAGTCACTATATGCT 416  
 Db 108 ProHisTyrLeuValProLysAspLeuArgAsnLysLysGlyLysIleIleTyrValGly 127  
 Qy 417 CGCAACCCAGGATCTGGTGGTCTTATATATGATTCACCGCTCTCTCGGACCATG 476  
 Db 128 ArgAsnProLysAspAlaAlaValSerTyrTyrHisPheTyrAsnValIleValArgLeu 147

QY 477 AGCTACCGAGGACCTTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTGGGCTAC 536  
 DB 148 LysGlnValAsnAspTrpGluSerPheLeuAspArgTyrLeuThrGlyGluValLeuGly 167  
 QY 537 GGCTCTCTGGTTTGACGACGTCGAGAGTTCTGGGAGCACCCGATCGACTCGAACCTGCTT 596  
 DB 168 GlySerTrpPheAspHisValLysGlyTyrTrpThrHisGlnGluAspPheAsnIleLeu 187  
 QY 597 TTTCTCAAGTATGAAGACATCATCGGACCTGGTGGACGATGGTGGAGCAGCTGGCCAGA 656  
 DB 188 PheValThrTyrGluGluMetLysLysAspLeuArgSerAlaValLeuLysIleCysLys 207  
 QY 657 TTCTCTGGGGGTGCTCTGTGACCAAGCCCGAGTGGAGCCCTGACCGAG----- 704  
 DB 208 PheValGluLysGluLeuAsnGluGlnValAspThrIleValGluLysAlaThrPhe 227  
 QY 705 -----CACTGCCACGACTGGTGGACCGAGTGGTGGACCGAGTGGCTGC 734  
 DB 228 LysAsnMetLysHisAspProLeuAlaAsnTyrThrAsnValSerThrAspHisLeuAsp 247  
 QY 735 AACGCTGAGGCCCTGCCGCGGGGAAAGAGTTGGGCTGTGGAAGGACATCTTCACC 794  
 DB 248 MetLysAsnGlyThrPheLeuArgGlyThrValGlyAspTrpLysLeuMetThr 267  
 QY 795 GTCTCCATGAATGAGAGTTTGTACTGTGTATAAACAGAGATGGGAAAGTGTGACCTC 854  
 DB 268 ValAlaGlnAsnGluLysPheAspLysIleTyrSerGluLysMetLysGlyValProIle 287  
 QY 855 ACGTTT 860  
 DB 288 AsnPhe 289  
 RESULT 88  
 ID ST1E1\_CAVPO STANDARD; PRT; 296 AA.  
 AC P49887;  
 DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.  
 DT 01-OCT-1996, sequence version 1.  
 DT 07-FEB-2006, entry version 43.  
 DE Estrogen sulfotransferase (EC 2.8.2.4) (Sulfotransferase, estrogen-  
 DE preferring) (ST1E3).  
 GN Name=SUT1E1; Synonyms=STE;  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 OC Hystricognathi; Caviidae; Cavia.  
 OX NCBI\_TaxID=10141;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [MRNA], AND PARTIAL PROTEIN SEQUENCE.  
 RC STRAIN=NIH 2; TISSUE=Adrenal cortex;  
 RX MEDLINE=93024479; PubMed=1406700; DOI=10.1210/me.6.8.1216;  
 RA Oeda T., Lee Y.C., Driscoll W.J., Chen H.-C., Strott C.A.;  
 RT "Molecular cloning and expression of a full-length complementary DNA  
 RT encoding the guinea pig adrenocortical estrogen sulfotransferase.";  
 RL Mol. Endocrinol. 6:1216-1226(1992).  
 RN [2]  
 RP PAPS-BINDING SITE, AND MUTAGENESIS.  
 RX MEDLINE=95071453; PubMed=7980593;  
 RA Komatsu K., Driscoll W.J., Koh Y., Strott C.A.;  
 RT "A P-loop related motif (GxxGxxK) highly conserved in  
 RT sulfotransferases is required for binding the activated sulfate  
 RT donor.";  
 RL Biochem. Biophys. Res. Commun. 204:1178-1185(1994).  
 CC -!- FUNCTION: Sulfation of estradiol and estrone. May control the  
 CC level of the estrogen receptor by sulfurylating free estradiol.  
 CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + estrone =  
 CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.  
 CC -!- SUBUNIT: Homodimer.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasm.  
 CC -!- TISSUE SPECIFICITY: Adrenal gland and much less in liver. Uterine  
 CC STE is detectable only during pregnancy.  
 CC -!- PTM: The N-terminus is blocked.

CC -1- SIMILARITY: Belongs to the sulfotransferase family.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; U09552; AAA18495.1; -; mRNA.  
 DR PIR; A44011; A44011.  
 DR HSSP; P49888; 1G3M.  
 DR SMK; P49887; 5-234.  
 DR InterPro; IPR000863; Sulfotransferase.  
 DR Pfam; PF00685; Sulfotransferase\_1; 1.  
 DR ProDom; PD001218; Sulfotransferase; 1.  
 KW Direct protein sequencing; Lipid-binding; Steroid-binding;  
 KW Transferrase.  
 FT CHAIN 1 296 Estrogen sulfotransferase.  
 FT /FTID=PRO\_0000085152.  
 FT NP\_BIND 49 54 PAPS (By similarity).  
 FT NP\_BIND 131 139 PAPS (By similarity).  
 FT NP\_BIND 194 230 PAPS (By similarity).  
 FT NP\_BIND 258 260 PAPS (By similarity).  
 FT ACT\_SITE 109 109 Proton acceptor (By similarity).  
 FT MUTAGEN 260 260 G->A: Loss of activity.  
 FT MUTAGEN 263 263 G->A: Loss of activity.  
 FT MUTAGEN 266 266 K->A: Loss of activity.  
 SQ SEQUENCE 296 AA; 35245 MW; 019D92895E29901D CRC64;  
 Alignment Scores:  
 Pred. No.: 4.59e-20 Length: 296  
 Score: 408.50 Matches: 94  
 Percent Similarity: 52.4% Conservative: 50  
 Best Local Similarity: 34.2% Mismatches: 112  
 Query Match: 9.3% Indels: 19  
 DB: 1 Gaps: 4  
 US-10-768-158-1 (1-2419) x ST1E1\_CAVPO (1-296)  
 QY 72 AAGTCTTC---GAGTTCATCGCTCGCGCTCGCGCTTCGCGCGGAAGATGAG 128  
 DB 11 GlutyrPheAspGluPheArgGlyIleLeuLeuTyrLysGlnPheIleLysTyrTrpAsp 30  
 QY 129 GAGATCGCAACTTCCTCCGCTGGCGCCGACGACGTGTGGATCGTCACCTACCCCAAGTCC 188  
 DB 31 AsnValGluAlaPheGlnAlaArgProAspAspLeuValIleAlaAlaTyrProLysSer 50  
 QY 189 GGCACAGCTTCGACGAGGTGTCTACTTGTGTGAGCCGCGCTGACCCCATGAG 248  
 DB 51 GlyThrThrTrpIleSerGluValValCysMetIleTyrAlaGluGlyAspValLysLys 70  
 QY 249 ATCGCTTGTATGAACATGACGACGACGTCCCGCTCGGTGAGTACCCACAGCCG----- 302  
 DB 71 CysArgGlnAspAlaIlePheAsnArgValProPheLeuGluCysArgAsnAspLysMet 90  
 QY 303 -----GGCTGGACATCATCAAGAACTGACCTTCCTCCGCTCATCAAGAGCCACCTG 356  
 DB 91 MetAsnGlyValLysGlnLeuGluMetAsnSerProArgIleIleLysThrHisLeu 110  
 QY 357 CCCTACCGCTTCTCCCTCTGACCTCCCAATGAGACTCCAAAGTATCATATATGCT 416  
 DB 111 ProProArgLeuProAlaSerPheTrpGluLysArgCysLeuMetIleCysLeu 130  
 QY 417 GCGACCCCAAGGATCTGTGTGTCTATTATCATGTTCCACCGCTCTCTCGGACCATG 476  
 DB 131 ArgAsnAlaLysAspValAlaValSerTyrTyrPhePheLeuMetValAlaHis 150  
 QY 477 AGCTACCGAGGACCTTTCAAGAAATCTCGCGGAGTTTATGAATGATGAAGCTGGCTAC 536  
 DB 151 ProAspProGlySerPheProGluPheValGluLysPheMetGlnGlnValProTyr 170  
 QY 537 GCCTCTGTTGACGACGTGAGGAGTCTGGGACCCGACCGATCGACTCGACCTGCTT 596  
 DB 171 GlySerTrpTyrAspHisValLysSerTrpTrpTrpGluLysSerThrAspProArgIleLeu 190  
 QY 597 TTCTCAAGTATGAAGACATGATCGGACCTCGGACCTGGTGACGATCGTGGAGCAGCTGGCCAGA 656

Db 191 PheIlePheTyrGluAspMetLysGluAspIleArgLysGluValLeuLysLeuHis 210  
 QY 657 TTCTTGGGGTGTCTGTGACAGCCAGCTGGAAGCCCTGAGCGACATGCGCAC--- 713  
 Db 211 PheLeuGlyArgLysProSerGluGluLeuValAspLysIleIleLysHisThrSerPhe 230  
 QY 714 -----CAGCTGGTGGACCACTGCTGCAAC 737  
 Db 231 GlnGluMetLysAsnAsnProSerThrAsnTyrThrMetLeuProGluGluLeuMetAsn 250  
 QY 738 GCTGAGGCCCTGCCC---GTGGCGCGGGAAGACTGTGGCTGTGGAAAGACATCTTCACC 794  
 Db 251 GlnLysValSerProPheMetArgLysGlyIleSerGlyAspTrpLysAsnHisPheThr 270  
 QY 795 GTCTCCATGAATGAGAGTTTGACTTGGTGTATATAACAGAGATG 839  
 Db 271 ValAlaLeuAsnGluSerPheAspLysHisTyrGlnGlnMet 285  
 RESULT 89  
 O70262 MOUSE PRELIMINARY; PRT; 304 AA.  
 AC O70262;  
 DT 01-AUG-1998, integrated into UniProtKB/TrEMBL.  
 DT 01-AUG-1998, sequence version 1.  
 DT 07-FEB-2006, entry version 26.  
 DE Phenol sulfotransferase.  
 GN Name=Sult1c1; Synonyms=Sult1a2;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=ddv; TISSUE=Olfactory;  
 RX MEDLINE=96208491; PubMed=8641270;  
 RA Miyawaki A., Homma H., Tamura H., Matsui M., Mikoshiba K.;  
 RT "Zonal distribution of sulfotransferase for phenol in olfactory  
 sustentacular cells.";  
 RL EMBO J. 15:2050-2055(1996).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=ddv; TISSUE=Olfactory;  
 RX MEDLINE=96228237; PubMed=9560327;  
 RA Tamura H.O., Harada Y., Miyawaki A., Mikoshiba K., Matsui M.;  
 RT "Molecular cloning and expression of a cDNA encoding an olfactory-  
 specific mouse phenol sulphotransferase.";  
 RL Biochem. J. 331:953-958(1998).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=ddv; TISSUE=Olfactory;  
 RA Tamura H.-O., Harada Y., Miyawaki A., Mikoshiba K., Matsui M.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AF033653; AAC17740.1; -; mRNA.  
 DR HSSP; P50224; 1CJM.  
 DR Ensembl; ENSMUSG00000023943; Mus musculus.  
 DR MGI; MGI-102928; Sult1c1.  
 DR GO; GO:0005737; Cytoplasm; IDA.  
 DR GO; GO:0004062; F-aryl sulfotransferase activity; IDA.  
 DR GO; GO:0006790; P:sulfur metabolism; IDA.  
 DR InterPro; IPR000863; Sulfotransferase.  
 DR Pfam; PF00685; Sulfotransferase\_1; 1.  
 DR ProDom; PD001218; Sulfotransferase; 1.  
 KW Transferrase.  
 SQ SEQUENCE 304 AA; 35797 MW; 30857261A1462BE4 CRC64;  
 Alignment Scores:  
 Pred. No.: 4.62e-20 Length: 304





```

QY 375 TCTGACCTCCAAATGAGAGCTCCAAAGCTCATATATGGCTGCGCAACCCCAAGGATCTG 434
Db 130 ProAlaPheTrpGluAsnLysCysLysThrIleTyrValAlaArgAsnAlaLysAspAsn 149
QY 435 GTGGTGCTTATTATCACTGCTCCACCGCTCTCTCGGACCATGAGTACTCCGAGGACACCTTT 494
Db 150 LeuValSerTyrTyrPhePheAspCysMetAsnMetThrGlnProGluProGlyThrMet 169
QY 495 CAAGAATTTCTGCGGAGGTTTATGAATGATGAAGCTGGGCTACGGCTCTGTTTGAAGCAG 554
Db 170 GluGluTyrIleHisLysPheMetArgGlyGluLeuSerTrpGlySerTrpTyrAspHis 189
QY 555 GTGAGGAGTCTGGGACACCGCATGACGACGACGACGACGACGACGACGACGACGACGACG 614
Db 190 ValLysGlyTyrTrpLysGluLysAspAsnLysAsnIleLeuTyrLeuPheTyrGluAsp 209
QY 615 ATGCATCGGACCTGGTACGATGGTGGAGCAGCTGGCAGATCTCTGGGGTGTCCTGT 674
Db 210 MetLysGluAsnProArgGluValGluArgIleMetArgTyrLeuAspValSerVal 229
QY 675 GACAAAG-----GCCAGCTGGAAGCCCTGACGAGGACGACGACGACGACGACGACG 725
Db 230 SerAspGluValIleSerLysIleValGluLeuThrSerPheGluLysMetLysGluAsn 249
QY 726 CAGTGCTGCAACGCTGAGCGCTGCC----- 752
Db 250 ProMetAlaAsnTyrThrCysValProAlaProValPheAspHisSerLysSerProPhe 269
QY 753 GTGGCGCGGGAAGAGTGTGGGCTGTGGAGGACATCTTCACCGTCTCCCATGAATGAGAAG 812
Db 270 MetArgLysGlyLysValGlyAspTrpArgAsnTyrPheThrProGluGlnGluLysMet 289
QY 813 TTGTGCTGGTGTTATAACAGAGATGGGAAGTGTGACCTCACGCTTT 860
Db 290 PheGluGluAspTyrLysGluGlnMetLysAspValAspIleProPhe 305

RESULT 91
QSVHH0 DROME
ID QSVHH0 DROME PRELIMINARY; PRT; 316 AA.
AC QSVHH0;
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 27.
DE 016733-PA.
GN ORFNames=CG16733, Dmel_CG16733;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
[1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Sutton G.G., Worland J.R., Richards S., Ashburner M., Henderson S.N.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster C., Gabrielson A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöckle A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

```

```

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E.R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Svirskaas R., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Wang Z.-Y., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195 (2000).
[2]
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RL melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).
[3]
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskaas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RL a genomics perspective."
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20 (2002).
[4]
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Mira S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RL systematic review."
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22 (2002).
[5]
RN NUCLEOTIDE SEQUENCE.
RG Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskaas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[6]
RN NUCLEOTIDE SEQUENCE.
RG FlyBase;
RL Submitted (JAN-2006) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC EMBL; AE003682; AAF54344.1; -; Genomic_DNA.
DR HSSP; P49888; IHY3.
DR FlyBase; FBgn0037665; CG16733.
DR GO; GO:0008146; F:sulfotransferase activity; IBA.
DR InterPro; IPR000863; Sulfotransferase.

```







```
QY 852 CTCACGTTTGACTTT 866
DB 287 LeuProIleArgPhe 291

RESULT 95
Q6IM15_HUMAN
AC Q6IM15 PRELIMINARY; PRT; 304 AA.
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 10.
DE SULT1C3 splice variant a.
GN Name=SULT1C3;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed:14767822; DOI=10.1038/sj.tpj.6500223;
RA Freimuch R.R., Wiepert M., Chute C.G., Wieben E.D., Weinshilboum R.M.;
RT "Human cytosolic sulfotransferase database mining: identification of
  seven novel genes and pseudogenes.";
RL Pharmacogenomics J. 4:54-65(2004).
CC -!- MISCELLANEOUS. The sequence shown here is derived from an
  EMBL/GenBank/DBJ third party annotation (TPA) entry.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
EMBL; BK001432; DAA01770.1; -; Genomic_DNA.
DR HSSP; P49891; 1AOY
DR GO; GO:0008146; F: sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer_1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
SQ SEQUENCE 304 AA; 35674 MW; 398CB6FF9792FFD9 CRC64;

Alignment Scores:
Pred. No.: 1.57e-19 Length: 304
Score: 401.00 Matches: 93
Percent Similarity: 51.7% Conservative: 62
Best Local Similarity: 31.0% Mismatches: 125
Query Match: 9.1% Indels: 20
DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x Q6IM15_HUMAN (1-304)
QY 21 ATGGCGGAGAGCGCGGAGAGACCCCGGAGCCCGGCGGAGTTCGAG---AGCAAGTAC 77
DB 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
QY 78 TTCGAGTTCATGGCGTGGCGGCTGCCGCCCTTCTCCCGGGAAGATGAGGAGATCGCC 137
DB 21 MetGluValAspGlyValProThrLeuLeuSerLysGluTrpTrpGluLysValCys 40
QY 138 AACTTCCCGTCCGCGGAGCGACGCTGGATCGTCACCTACCCCAAGTCCGGCACCAGC 197
DB 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTyrProLysSerGlyThrThr 60
QY 198 TTGCTTCGAGGAGGTGGTCTACTTGTGTGAGCCAGCGGCGGTGACCCGATGAGATCGGCTG 257
DB 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
QY 258 ATGAATCGACGAGAGAGCTCCGCTCTCGAG-----TACCCA-----CAGCGC 302
DB 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysLysGluLysPro 100
QY 303 GGCCTGGACATCAAGGAACCTGACCTCCCGCCCTCATCAAGGACCACTGCGCTAC 362
DB 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120

363 CGTTTCTGCTCTGACCTCCCAATGAGAGCTCAAGGTCACTATATATGCTGCGCAAC 422
DB 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
QY 423 CCAAGGATCTGGTGGTCTTATATATCAGTTCCACCGCTCTCTGCGGACCATGAGTAC 482
DB 141 ProLysAspCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAsp 160
QY 483 CGAGGCACCTTTCAAGAAATCTCGCGGAGGTTTATGATGATAAGCTGGGTACGGCTCC 542
DB 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlyGlySer 180
QY 543 TGGTTTGAGCACGTCGAGGAGTCTGGGAGCACCGCATGCGACTCGCAACGTCTTTTCTC 602
DB 181 TrpPheAsnHisValLysGlyTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200
QY 603 AAGTATGAAGACATGTCATCGGACCTGGTGACGATGGTGGAGCAGCTGGCCAGATTCTCTG 662
DB 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
QY 663 GGGGTGCTCTGTGACAAAGCCACGCTGGAAGCCCTGACGGAGCACTGCCACCAGCTGGTG 722
DB 221 GluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
QY 723 -----GACCAGTGTGCAACGCTGAGCGCTGCGCC----- 752
DB 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
QY 753 -----GTGGCGCGGGAAGAGTGGCTGTGGAAGAGCACTTTCACCGTCTCC 800
DB 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
QY 801 ATGAATGAGAAAGTTTGACTTGGTGTATAAACAGAGATGGGAAAGTGAGCCTCACGTTT 860
DB 281 LeuAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300

RESULT 96
ST3A1_RABIT
ID ST3A1_RABIT STANDARD; PRT; 301 AA.
AC O46640;
DT 07-FEB-2006, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-1998, sequence version 1.
DT 21-FEB-2006, entry version 26.
DE Amine sulfotransferase [EC 2.8.2.3] (Sulfotransferase 3A1) (AST-RB1).
GN Name=SULT3A1; Synonyms=ST3A1;
OS Oryctolagus cuniculus (Rabbit);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], PROTEIN SEQUENCE OF 29-66; 160-176;
RP 200-206; 254-260 AND 286-301, FUNCTION, TISSUE SPECIFICITY, AND
RP SUBCELLULAR LOCATION.
RC STRAIN=New Zealand white; TISSUE=Liver;
RX MEDLINE=98207028; PubMed=9538231;
RA Yoshinari K., Nagata K., Ogino M., Fujita K., Shiraga T., Iwasaki K.,
  Hata T., Yamazoe Y.;
RT "Molecular cloning and expression of an amine sulfotransferase cDNA: a
  new gene family of cytosolic sulfotransferases in mammals.";
RL J. Biochem. 123:479-486(1998).
CC -!- FUNCTION: Catalyzes N-sulfation of amines (PTHF, aniline, 4-
  chloroaniline, 2-naphthylamine).
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylyl sulfate + an amine =
  adenosine 3',5'-bisphosphate + a sulfamate.
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- TISSUE SPECIFICITY: Expressed in male liver.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
EMBL; D86219; BAA24994.1; -; mRNA.
```

Qy	855	ACGTTTCACTTTATTATTAATAACAGAACACAACCTCATGCT	900
Db	285	LysPheilleTrpAspLeu-AsnAspGluAlaAsnSerAsnHisSer	299
RESULT 97			
Q53SG4_HUMAN			
ID	Q53SG4_HUMAN	PRELIMINARY; PRT;	307 AA.
AC	Q53SG4;		
DT	24-MAY-2005,	integrated into UniProtKB/TrEMBL.	
DT	24-MAY-2005,	sequence version 1.	
DE	07-FEB-2006,	entry version 3.	
DE	Hypothetical protein SULNIC1.		
GN	Name=SULNIC1;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;		
OC	Homo.		
OX	NCBITaxID=9606;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Kalicki J., Drone K., Belter E.;		
RT	"The sequence of Homo sapiens BAC clone RP11-443K8.";		
RL	Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Waterston R.;		
RL	Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.		
RN	[3]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Waterston R.;		
RL	Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.		
RN	[4]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Wilson R.K.;		
RL	Submitted (APR-2005) to the EMBL/GenBank/DDBJ databases.		
CC			
CC	Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms		
CC	Distributed under the Creative Commons Attribution-NoDerivs License		
CC	-----		
DR	SMR; AC019100; AAV14790.1; -; Genomic_DNA.		
DR	ENR; Q53SG4; 12-307.		
DR	Ensembl; ENSG00000198203; Homo sapiens.		
DR	GO; GO:0008146; F:sulfotransferase activity; IEA.		
DR	InterPro; IPR000863; Sulfotransferase.		
DR	Pfam; PF00685; Sulfotransferase 1; 1.		
DR	ProDom; PD001218; Sulfotransferase; 1.		
KW	Hypothetical protein.		
SQ	SEQUENCE 307 AA; 35868 MW; DB83191FA88BFD1 CRC64;		
Alignment Scores:			
Pred. No.:	3,88e-19	Length:	307
Score:	395.50	Matches:	89
Percent Similarity:	49.2%	Conservative:	56
Best Local Similarity:	30.2%	Mismatches:	115
Query Match:	9.0%	Indels:	35
DB:	2	Gaps:	5
US-10-768-158-1 (1-2419) x Q53SG4_HUMAN (1-307)			
Qy	72	AAGTACTTCGAGTTCCATGGCGGTGCCTGCCGCCCTTCGCCCGCGGAAGATGGAGGAG	1
Db	12	LysLeuLysGluValGluGlyThrLeuLeuGlnProAlaThrValAspAsnTrpSerGln	3
Qy	132	ATGCCCAACTTCCCGTGCGCGCAGCGCTGTGGATCGTCACCTACCCTACCCAGTCCGGC	1
Db	32	IleGlnSerPheGluAlaLysProAspAspLeuLeulleCysThyrPrPoLYsAlaGly	5
Qy	192	ACCAGCTTGCTGCAGGAGGTGGTCTACTTGGTGAGCCAGCGCGGTGACCCCGATGATC	2
Db	52	ThrThrTrpIleGlnGluIleValAspMetIleGlnAsnGlyAspValGluLyseCys	7
Qy	252	GGCTTGATGAACATCGACGACGCTCCCGGTCTCTGGAGTAC-----CCACAGCCG	3

```
Db 72 GlnArgAlaIleIleGlnHisArgHisProPheIleGluTrpAlaAArgProGlnPro 91
Qy 303 -----GGCCTGGACATCATCAAGAACTGACC 329
Db 92 SerGluThrGlyPheHisValAlaGlnGlyLeuLysLeuLeuSerSerSerAsn 111
Qy 330 TCTCCCGCCTCATCAAGAGCCACCTGCCCTACCGC---TTTCTGCCCTCTGACCTCCAC 386
Db 112 ProProAlaSerThrSerGlnSerAlaLysIleThrAspLeuLeuProProSerPheTrp 131
Qy 387 AATGAGACTCCAAGGTCTATATGCTCGCAACCCCAAGGATCTGGTGTCTTAT 446
Db 132 GluAenAenCysLysPheLeuTrpValAlaAargAenAlaLysAAspCysMetValSerTyr 151
Qy 447 TATCAGTTCCACCGCTCTCTGGGACCATGAGCTACCGAGGACCTTTCAGAATTCTGC 506
Db 152 TyrHisPheGlnAArgMetAenHisMetLeuProAspProGlyThrTrpGluGluTyrPhe 171
Qy 507 CGGAGTTTATGAATGATAAGCTGGCTACCGCTCTCTGTTTGAGCAGCTGCAGGAGTTC 566
Db 172 GluThrPheIleAenGlyLysValValTrpGlySerTrpPheAspHisValLysGlyTrp 191
Qy 567 TGGGAGCACCGCATGAGCTCGAAGCTGCTTTTCTCAAGTATGAAGACATCATCGGAC 626
Db 192 TrpGluMetLysAAspArgHisGlnIleLeuPheLeuPheTyrGluAAspIleLysAArgAsp 211
Qy 627 CTGGTGACATGTTGGAGCAGCTGCCAGATTCCTGGGGGTGCTCTGTGACAAAGCCGAC 686
Db 212 ProLysHisGluIleAArgLysValMetGlnPheMetGlyLysLysValAAspGluThrVal 231
Qy 687 CTGGAAGCCCTGACGGAGCAC----- 707
Db 232 LeuAAspLysIleValGlnGluThrSerPheGluLysMetLysGluAAsnProMetThrAsn 251
Qy 708 -----TGCCACACAGCTGTGGACCATGCTCTGCAACGCTGAGCCCTGCCCGTG 755
Db 252 ArgSerThrValSerLysSerIleLeuAAspGlnSerIleSerSerPhe-----Met 268
Qy 756 GCCCGGGAAGTTGGCTGTGGAGACATCTTCACCGTCTCCATCAATGAGAAGTTT 815
Db 269 ArgLysGlyThrValGlyAspTrpLysAenHisPheThrValAlaGlnAAsnGluAArgPhe 288
Qy 816 GACTTGGTGTTAAACAGAGAAGTGGAAAGTGTGACCTCACGTTT 860
Db 289 AspGluIleTyrArgArgLysMetGluGlyThrSerIleAAsnPhe 303

RESULT 98
ID Q6DJ68_XENTR PRELIMINARY; PRT; 304 AA.
AC Q6DJ68_
DT 16-AUG-2004, integrated into UniProtKB/TrEMBL.
DT 16-AUG-2004, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE MGC88979 protein.
GN Name=MGC88979;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole body;
RX MEDLINE=2388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ussin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguercio N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
```

```
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Paney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blackley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krawinski M.I., Skalska U., Smallos D.E.,
RA Snerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole body;
RA Klein S., Gerhardt D.S.; to the EMBL/GenBank/DBJ databases.
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DB EMBL; BC075315; AAH75315.1; -; mRNA.
DR Ensembl; ENSXETG0000009267; Xenopus tropicalis.
DR GO; GO:0008146; F:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase 1; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
SQ SEQUENCE 304 AA; 35586 MW; F73BE91748AE2CFA CRC64;

Alignment Scores:
Pred. No.: 5,36e-19 Length: 304
Score: 393.50 Matches: 88
Percent Similarity: 54.8% Conservative: 56
Best Local Similarity: 33.5% Mismatches: 100
Query Match: 8.9% Indels: 19
DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x Q6DJ68_XENTR (1-304)
Qy 129 GAGATCGCAACTTCCCGGTGCGGCGCCAGCAGCTGGATCGTCACCTACCCCAAGTCC 188
Db 38 AspIleTrpAenPheArgAlaHisProGluAAspValLeuLeuAAspSerTyrProLysAla 57
Qy 189 GGCACCAAGCTGTCTGACGAGGTGTCTACTTGTGTGAGCAGCGGCGCTGACCCCGATGAG 248
Db 58 GlyIleThrTrpMetGlnGluValValAspMetIleTyrGlnGluGlyAspThrAsnLys 77
Qy 249 ATCGGCTTGATGAACATCGACGACGACTCCCGTCTCTGGAGTAC-----CCACAGCG 302
Db 78 CysLeuArgAlaProThrTyrAspArgHisProPheLeuLeuGluAlaValProLysPro 97
Qy 303 -----GGCCTGGACATCATCAAGAACTGACCTCTCCCGCCTCATCAAGAGCCAC 353
Db 98 ValProSerGlyLeuGlnLeuAAspMetGluProProArgValLeuLysThrHis 117
Qy 354 CTGCGCTACCCCTTCTGCGCTCTGACCTCCACCAATGGAGACTCCAAAGTCTATATATG 413
Db 118 LeuProIleGlnLeuIleProProSerPheTrpLysGlnAAspCysLysValIleTyrVal 137
Qy 414 GCTCGCAACCCCAAGGATCTGGTGTGTCTTATATCATAGTTCCACCGCTCTCTCGGAGCC 473
Db 138 AlaArgAenAlaLysAAspSerLeuValSerTyrPheHisPheGlnArgMetThrLysGly 157
Qy 474 ATGAGTACCCAGGACCTTTCAAGAATTCTGCGGAGGTTTATCAATGATAAGTGGC 533
Db 158 LeuProAAspProGlyThrTrpAAspTyrPheMetAlaPheLeuSerGlyThrLeuPro 177
Qy 534 TACGGCTCTCTGGTTGACACGCTGAGGAGTTCTGGGAGCACCGCATCGACTCGAACCTG 593
Db 178 TrpGlySerTrpPheAspHisValAsnGlyTrpTrpLysAlaLysAAspHisArgVal 197
Qy 594 CTTTTCCTCAAGTATGAAGACATCATCGGACCTCGGACCTGTGTGATGGTGGAGAGCTGCC 653
Db 198 LeuTyrValPheTyrGluAAspMetLysAAspLeuArgLeuGluIleGlnArgValGlu 217
```



QY 654 AGATTCTCTGGGGTGTCTCTGTGACAAAGCCAGCTGGAAGCCCTGACGGAGCAC---TGC 710  
 |||||  
 Db 218 SerPheLeuAspLysAspLeuProGluGluValLeuGluLysIleCysGlnHisThrThr 237  
 |||||  
 QY 711 CACACAGCTGGTGACACAG-----TGCTGCACAGCTGAGGCCCTCGCC----- 752  
 |||||  
 Db 238 PheGlnAlaMetLysGluAsnProMetAlaAsnTyrThrThrMetProThrThrValMet 257  
 |||||  
 QY 753 -----GTGGCGCGGGAAGAGTTGGGCTGTGGAAGGACACATCTTC 791  
 |||||  
 Db 258 AspGlnSerValSerProPheMetArgLysGlyIleValGlyAspTrpLysAsnHisPhe 277  
 |||||  
 QY 792 ACCGTCTCCATGATGAGAGTTGACCTGGTGTTATTAACAGAGATGGAAGTGTGAC 851  
 |||||  
 Db 278 LeuValAlaGlnAsnGluLeuPheAspTrpGluTyrLysArgMetAspGlyThrGly 297  
 |||||  
 QY 852 CTCACGTTT 860  
 |||||  
 Db 298 LeuAspPhe 300  
 |||||

RESULT 99  
 ID Q7ZYHO\_XENLA PRELIMINARY; PRT; 304 AA.  
 AC Q7ZYHO;  
 DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.  
 DT 01-JUN-2003, sequence version 1.  
 DT 07-FEB-2006, entry version 17.  
 DE Sult1c1-prov protein.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;  
 OC Xenopodinae; Xenopus; Xenopus.  
 NCBI\_TaxID=8355;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative";  
 RL Dev. Dyn. 225:384-391 (2002).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Embryo;  
 RX Klein S., Strausberg R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; BC043790; AAH43790.1; -, mRNA.  
 DR HSSP; P50224; 1CJM.  
 DR GO; GO:0008146; F:sulfotransferase activity; IEA.  
 DR InterPro; IPR000863; Sulfotransferase.  
 DR Pfam; PF00685; Sulfotransfer\_1; 1.  
 DR ProDom; PD001218; Sulfotransferase; 1.  
 SQ SEQUENCE 304 AA; 35108 MW; 0D4C0C2B017C5692 CRC64;  
 Alignment Scores: 5.36e-19 Length: 304  
 Pred. No.: 393.50 Matches: 95  
 Score: 393.50  
 Percent Similarity: 50.2% Conservative: 49  
 Best Local Similarity: 33.1% Mismatches: 120  
 Query Match: 8.9% Indels: 23  
 DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x Q7ZYHO\_XENLA (1-304)

QY 63 TTCGAGAGCAAGTACTTCGAGTTCATGGCTGGGGTGGCGCCCTTCTGCGCGGGAAG 122  
 |||||  
 Db 16 PheGlnValThrMetGlyHisIleGluGlyValProLeuProValThrThrCysAspMet 35  
 |||||

QY 123 ATGGAGGAGATCGCAACTTCCCGTGGCGCCCGACGACGTGTGGATCGTACCTACCC 182  
 |||||  
 Db 36 TrpAspThrIleTyrAsnPheGlnAlaArgLysAspAlaLeuIleAlaThrThrPro 55  
 |||||

QY 183 AAGTCCCGCACCACTGCTGCAGAGGTGTCTACTTGTGTGAGCCAGGCGGTGAC--- 239  
 |||||  
 Db 56 LysAlaGlyThrTrpMetGlnGluIleValAspLeuIleLeuGlnGluGlyAspVal 75  
 |||||

QY 240 -----CCCGATGAGATCGGCTTGATGAACATCGACGACGAGCTCCCG 281  
 |||||  
 Db 76 GlnLysSerMetArgAlaProCysTyrIleLysValProPheIleAsp---LeuValPro 94  
 |||||

QY 282 GTCCTGGAGTACCACAGCGCGGCTGGACATCATCAAGAACTGACCTCTCCCGCCTC 341  
 |||||  
 Db 95 ProLysProMetPro---ProGlyValAlaLeuAlaGlnThrMetAsnSerProArgIle 113  
 |||||

QY 342 ATCAAGACCACTCGCCCTACCGCTTCTGCGCTTGCACCTCCACATGGAGACTCCAAG 401  
 |||||  
 Db 114 LeuLysThrHisLeuProIleAsnLeuLeuProSerPheTrpGluLysAsnThrLys 133  
 |||||

QY 402 GTCATCTATATGGCTCGCAACCCCAAGATCTGGTGTCTTATTATCAGTTCCACCGC 461  
 |||||  
 Db 134 ValValTyrValAlaAlaGlnAlaLysAspSerMetValSerTyrTyrPheHisLys 153  
 |||||

QY 462 TCTCTCGGACCATGAGTACCGAGCACCTTTTCAAGAATTTCTCGCGAGGTTTATGAAT 521  
 |||||  
 Db 154 MetAsnLysPheLeuProAspSerGlyThrLeuAspAsnPhePheSerGluPheLeuSer 173  
 |||||

QY 522 GATAAGCTGGGTACGGCTCTGGTTGACACCTGCGGAGGTTCTGGGACACCGCATG 581  
 |||||  
 Db 174 GlyAspValProTrpGlySerTrpPheAspAsnValLeuGlyTyrTrpTrpAlaLeuAsp 193  
 |||||

QY 582 GACTCGAAACGTGCTTTTCTCAAGATATCAAGACATCATCGGACCTGGTGCACGTGTG 641  
 |||||  
 Db 194 LysHisGlnIleLeuPheIlePheTyrGluAspMetIleGlnAspProMetArgGluIle 213  
 |||||

QY 642 GAGCAGCTGGCCAGATTCTGGGGGTGTCTGTGTGACAAAGGCCAGCTGGAAGCCCTGACG 701  
 |||||  
 Db 214 LysLysValMetThrPheLeuGlyLysAspLeuSerAspGluValLeuGluLysIleLys 233  
 |||||

QY 702 GAGCAGCTGCCAC-----CAGCTGGTGACACGTGTGCAACGCTGAGGCCCTGCC 752  
 |||||  
 Db 234 TyrHisThrSerPheGlnAlaMetLysGluAsnProMetThrAsnAsnSerThrValPro 253  
 |||||

QY 753 -----GTGGCGCGGGAAGAGTTGGGCTGGG 779  
 |||||

Db 254 LysThrIleMetAspGlnThrIleSerProPheIleArgLysGlyThrValGlyAspTrp 273  
 |||||





```

c 88      146      3.3      1039      2      T35878      hypothetical prote
c 89      145.5    3.3      251      1      PIHUPF      salivary proline-r
c 90      145      3.3      673      1      CGB06C      collagen alpha 1(I
c 91      145      3.3      1418      2      T4546C      collagen alpha 1(I
c 92      144.5    3.3      283      2      T19732      hypothetical prote
c 93      144.5    3.3      286      2      S3466S      collagen 1 - Caeno
c 94      144.5    3.3      296      2      A31219      collagen alpha 1(I
c 95      144      3.3      1487      1      CGHU6C      hypothetical prote
c 96      143.5    3.3      301      2      T21314      hypothetical prote
c 97      143.5    3.3      316      2      T20497      hypothetical prote
c 98      143.5    3.3      333      2      T5235S      collagen alpha 1(I
c 99      143.5    3.3      671      1      CGRT1S      collagen alpha 1(I
100      143.5    3.3      1106      2      JQ040S      hypothetical 119.5

RESULT 1
S28183
A:Title: Mus musculus (house mouse)
C:Species: Mus musculus (house mouse)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S28183
R:Kong, A.N.T.; Ma, M.; Tao, D.; Yang, L.
Biochim. Biophys. Acta 1171, 315-318, 1993
A:Title: Molecular cloning of cDNA encoding the phenol/aryl form of sulfotransferase (ms
A:Reference number: S28183, MUID:93144346; PMID:8424956
A:Accession: S28183
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-298 <KON>
A:Cross-references: UNIPROT:Q9R155; UNIPARC:UPI0000175825
A:Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 51-Tyr
C:Superfamily: alcohol sulfotransferase
C:Keywords: sulfotransferase

Alignment Scores:
Pred. No.:      Length:      298
Score:          495.00      Matches:      104
Percent Similarity: 53.9%      Conservative: 54
Best Local Similarity: 35.5%      Mismatches:  117
Query Match:     11.2%      Indels:       18
DB:              Gaps:       3

US-10-768-158-1 (1-2419) x S28183 (1-298)

QY      36      GCGGAGACCCCGACGACCCGCGGGAGTTCGAGAGCAAGTACTTCGAGTTCCTCCATGGCGTG 95
Db      2      AlaGlnAsnProSerAsnMetGluProLeuArgLysProLeuValProValLysGlylle 21
QY      96      CGGCTGCCGCCCTTCTTCGCCGCGGGAAGATCGAGGAGATCGCCAACTTCCCGTGGCGGCC 155
Db      22      ProLeuIleLysTyrPheAlaGluThrMetGluGlnLeuGlnAsnPheThrAlaTyrPro 41
QY      156      AGCAGCTGTGATCTGATCTACCTACCCCAAGTCGGGACGAGCTGCTGTCGAGGAGTGTC 215
Db      42      AspAlaValLeuIleSerThrTyrProLysSerGlyThrTrpMetSerGluIleMet 61
QY      216      TACTTGGTGACGCGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGACGAGCAG 275
Db      62      AspMetIleTyrGlnGlyLysLeuAspLysCysGlyArgAlaProValTyrAlaArg 81
QY      276      CTCCTCGGCTCTCGAGTACCCA-----CAGCCGGGCTGGACATCATCAAGGAA 323
Db      82      IleProPheLeuGluPheSerCysProGlyValProProGlyLeuGluThrLeuLysGlu 101
QY      324      CTGACCTCTCCCGCTCATCAAGAGCCAGCTCCCTACCGCTTTCGCGCTCTGACCTC 383
Db      102      ThrProAlaProArgIleIleLysThrHisLeuProLeuSerLeuLeuProGlnSerLeu 121
QY      384      CACAATGGAGACTCCCAAGGTTCATCTATATGGCTCGCAACCCCAAGGATCTGGTGGTGT 443
Db      122      LeuAspGlnLysIleLysValIleTyrValAlaArgAsnAlaLysAspValValSer 141

```

```

QY      444      TATTATCAGTTCCACCGCTCTCTCGGGACCATGAGTACCAGGACACCTTTCAAGAATTC 503
Db      142      TyrTyrAsnPheTyrLysMetAlaLysLeuHisProAspProGlyThrTrpGluSerPhe 161
QY      504      TGCCGGAGTTTATGAATGATAAGCTGGCTACGCTCCTGTTTGACACGCTGCAGGAG 563
Db      162      LeuGluAsnPheMetAspGlyLysValSerTyrGlySerTrpTyrGlnHisValysGlu 181
QY      564      TTCTGGGAGCACCGCATGAGCTCGAACGTCGTTTTTCTCAAGTATGAAGACATGCGG 623
Db      182      TrpTrpGluLeuArgArgThrHisProValLeuTyrLeuPheTyrGluAspMetLysGlu 201
QY      624      GACCTGGTGACGATGTGGGACGAGCTGCCAGATTCCTGGGGGTGCTCTGTGACAAAGCC 683
Db      202      AsnProLysArgGluIleLysLysIleLeuGluPheLeuGlyArgSerLeuProGluGlu 221
QY      684      CAGCTGGAAGCCCTGACGAGCACTGC-----CACCAGCTGGTGGACACAGTGCTGC 734
Db      222      ThrValAspLeuIleValHisThrSerPheLysLysMetLysGluAsnProMetAla 241
QY      735      AACGCTGAGGCCCTGCCCC-----GTGGGCCGG 761
Db      242      AsnTyrThrThrIleProThrGluValMetAspHisThrIleTyrProPheMetArgLys 261
QY      762      GGAAGAGTTGGCTGTGGAAGACATCTTCACCGTCTCCATGAATGAGAAGTTGACTTG 821
Db      262      GlyThrIleGlyAspTrpLysAsnThrPheThrValAlaGlnSerGluHisPheAspAla 281
QY      822      GTGTATTAACAGAGATCGGAAAGTGTGACCTCACGTTT 860
Db      282      HisTyrAlaLysLeuMetThrGlyCysAspPheThrPhe 294

RESULT 2
A55451
A:Title: aryl sulfotransferase (EC 2.8.2.1) HAST3 / estrogen sulfotransferase EST - human
N:Alternate names: gene STM protein; thermolabile phenol sulfotransferase
C:Species: Homo sapiens (man)
C:Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004
C:Accession: A55451; JN0713; JC2067; S52398; I38551; I57679; I52299; PC2031
R:Bernier, F.; Leblanc, G.; Labrie, F.; Luu-The, V.
J. Biol. Chem. 269, 28200-28205, 1994
A:Title: Structure of human estrogen and aryl sulfotransferase gene. Two mRNA species i-
A:Reference number: A55451; MUID:95050600; PMID:7961757
A:Accession: A55451
A:Molecule type: DNA
A:Residues: 1-295 <BER>
A:Cross-references: UNIPROT:P50224; UNIPARC:UPI0000046BED; GB:I34160
R:Zhu, X.; Veronese, M.E.; Bernard, C.C.A.; Sansom, L.N.; McManus, M.E.
Biochem. Biophys. Res. Commun. 195, 120-127, 1993
A:Title: Identification of two human brain aryl sulfotransferase cDNAs.
A:Reference number: JN0713; MUID:93371391; PMID:8363592
A:Accession: JN0713
A:Molecule type: mRNA
A:Residues: 1-295 <ZHU>
A:Cross-references: UNIPARC:UPI0000046BED; GB:L19956; NID:9306456; PIDN:AAA02943.1; PID
R:Wood, T.C.; Aksoy, I.A.; Aksoy, S.; Weinshilboum, R.M.
Biochem. Biophys. Res. Commun. 198, 1119-1127, 1994
A:Title: Human liver thermolabile phenol sulfotransferase: cDNA cloning, expression and
A:Reference number: JC2067; MUID:94161723; PMID:8117269
A:Accession: JC2067
A:Molecule type: mRNA
A:Residues: 1-295 <WOO>
A:Cross-references: UNIPARC:UPI0000046BED; GB:U08032; NID:9468256; PIDN:AAA17723.1; PID
A:Experimental source: liver
A:Experimental source: liver
R:Jones, A.L.; Hagen, M.; Coughtrie, M.W.H.; Roberts, R.C.; Glatt, H.R.
submitted to the EMBL Data Library, February 1995
A:Description: Human platelet phenol/sulfotransferases: cDNA cloning, stable expression
A:Reference number: S52398
A:Accession: S52398
A:Status: preliminary
A:Molecule type: mRNA

```

A:Residues: 1-295 <JON>  
A:Cross-references: UNIPARC:UPI0000046BED; EMBL:X84653; NID:G671533; PIDN:CAA59146.1; PI  
R:Aksoy, I.A.; Callen, D.F.; Apostolou, S.; Her, C.; Weinshilboum, R.M.  
Genomics 23, 275-277, 1994  
A:Title: Thermolabile phenol sulfotransferase gene (STM): Localization to human chromosome  
A:Reference number: A55725, MUID:95130098; PMID:7829089  
A:Accession: I38551  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 139-198 <RES>  
A:Cross-references: UNIPARC:UPI000016A097; EMBL:U08099; NID:G488284; PIDN:AAA82126.1; PI  
R:Bernier, F.; Lopez-Solache, I.; Labrie, F.; Luu-The, V.  
Mol. Cell. Endocrinol. 99, 11-15, 1994  
A:Title: Cloning and expression of cDNA encoding human placental estrogen sulfotransferase  
A:Reference number: I57679; MUID:94244847; PMID:8187952  
A:Accession: I57679  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-295 <RE2>  
A:Cross-references: UNIPARC:UPI0000046BED; GB:I25275; NID:G463124; PIDN:AAA36523.1; PID:  
R:Doolley, T.P.; Probst, P.; Munroe, P.B.; Mole, S.E.; Liu, Z.; Doggett, N.A.  
Biochem. Biophys. Res. Commun. 205, 1325-1332, 1994  
A:Title: Genomic organization and DNA sequence of the human catecholamine-sulfating phen  
A:Reference number: I52299; MUID:95100963; PMID:7802665  
A:Accession: I52299  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-295 <RE3>  
A:Cross-references: UNIPARC:UPI0000046BED; GB:S74971; NID:G833898  
C:Comment: This enzyme transfers a sulfonate moiety from 3'-phosphoadenosine-5-phosphosu  
C:Comment: Brain aryl sulfotransferase and placental estrogen sulfotransferase are trans  
C:Comment: differ only in their 5'-untranslation regions.  
C:Genetics:  
A:Gene: GDB:STE  
A:Cross-references: GDB:342104; OMIM:600043  
A:Map position: 16p11.2-16p11.2  
A:Introns: 50/1; 92/1; 124/3; 167/1; 198/3; 259/1  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: brain; placenta; sulfotransferase

Alignment Scores:  
Pred. No.: 1,098-27 Length: 295  
Score: 472.00 Matches: 100  
Percent Similarity: 54.5% Conservative: 52  
Best Local Similarity: 35.8% Mismatches: 101  
Query Match: 10.7% Indels: 26  
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x A55451 (1-295)

QY 90 GCGCTGCGGCTCGCGCCCTTCTGCGCGGGAAGATGGAGGATCGCCAACTTCCCGGTG 149  
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36

QY 150 CGGCCAGCAGCTGTGGATGCTACCTACCCCAAGTCGGACACGCTTGTCTGAGGAG 209  
Db 37 ArgProAspLeuLeuIleAsnThrTyrProLysSerGlyThrThrTrpValSerGln 56

QY 210 GTGGTCTACTTGGTGAGCGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269  
Db 57 IleLeuAspMetIleTyrGlnGlyAspLeuGlyLysCysAsnArgAlaProIleTyr 76

QY 270 GAGCAGCTCCGCTCGTGAGTACCCACAGCCG-----GGCCTGGACATCATC 317  
Db 77 ValArgValProPheLeuGluValAsnAspProGlyGluProSerGlyLeuGluThrLeu 96

QY 318 AAGGAAGTACCTCTCCCGCTCATCAAGAGCCACTCGCCTACCGCTTCTGCGCTCT 377  
Db 97 LysAspThrProProArgLeuIleLysSerHisLeuProLeuAlaLeuLeuProGln 116

QY 378 GACCTCCCAANTGGAGATCCCAAGTCTATGCTCGACACCCCAAGCATCTGGTG 437  
Db 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnProLysAspValAla 136

QY 438 GTGCTTTATTATCACTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGGACACTTTCAA 497  
Db 137 ValSerTyrTyrHisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAsp 156  
QY 498 GAAATCTCGCGAGGTTTATCAATGATGAAGCTGGCTACGGCTCCTGTTTGGAGCACGTG 557  
Db 157 SerPheLeuGluLysPheMetAlaGlyGluValSerTyrGlySerTyrPyrGlnHisVal 176  
QY 558 CAGGAGTTCTGGGAGCACCGCATGCAACGTGCTTTTTTCTCAAGTATGAAGACATG 617  
Db 177 GlnGluTyrTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196  
QY 618 CATCGGACCTGCTGACCATGCTGGAGCAGCTGGCCAGATTCCTGGGGTGTCTGTGAC 677  
Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216  
QY 678 AAGGCCAGCTGGAAGCCCTGACGAGCAGCTGC----- 710  
Db 217 GluGluThrMetAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236  
QY 711 -----CACCAGCTGGTGAGCAGCTGTGCAACCTGAGGCC 746  
Db 237 MetThrAsnTyrThrThrValProGlnGluLeuMetAspHisSerIleSer----- 253  
QY 747 CTGCCC---GTGCGCGCGGGAAGATTGGCTGTGGAAGGACATCTTACCGCTCTCCATG 803  
Db 254 ---ProPheMetArgLysGlyMetAlaGlyAspTyrLysThrPheThrValAlaGln 272  
QY 804 AATGAGAAGTTGACTTGGTGTATAAACAGAGAGTGGGAAAGTGACCTCAGCTTT 860  
Db 273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 3  
JC7282  
hydroxyarylamine sulfotransferase (EC 2.8.2.-) 2 - rat  
N:Alternate names: SULF1C2 sulfotransferase  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: JC7282  
R:Li, X.; Joehnk, C.; Hartmann, D.; Schestag, F.; Kroemer, W.; Gieselmann, V.  
Biochem. Biophys. Res. Commun. 272, 242-250, 2000  
A:Title: Enzymatic properties, tissue-specific expression, and lysosomal location of two  
A:Reference number: JC7282  
A:Contents: Kidney  
A:Accession: JC7282  
A:Molecule type: mRNA  
A:Residues: 1-296 <LIA>  
A:Cross-references: UNIPROT:Q9WUW8; UNIPARC:UPI00001360FC; GB:AJ238391  
C:Genetics:  
A:Gene: sulf1c2  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: kidney; sulfotransferase

Alignment Scores:  
Pred. No.: 2,68-27 Length: 296  
Score: 467.00 Matches: 100  
Percent Similarity: 55.4% Conservative: 58  
Best Local Similarity: 35.1% Mismatches: 105  
Query Match: 10.6% Indels: 22  
DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x JC7282 (1-296)

QY 66 GAGACAAGTACTTCCAGTTCATGCGCTGCGGCTCGCCCTTCTGCGCGGAAGATG 125  
Db 10 GlnThrLysLeuLysGluValAlaGlyIleProLeuGlnAlaProThrValAspAsnTrp 29

QY 126 GAGGAGATCGCAACTTCCCGTGGCGGCCAGCGCTGTGGATCGTCACTTACCCCAAG 185  
Db 30 SerGlnIleGlnThrPheLysAlaLysProAspLeuLeuIleCysThrTyrProLys 49

QY 186 TCCGACACAGCTTCTGCGAGGAGTGTCTACTTGGTGAGCCAGGCGCTGACCCCGAT 245

```
Db 50 SerGlyThrThrTrpIleGlnGluValAspMetIleGluGlnAsnGlyAspValGlu 69
QY 246 GAGATCGGCTTGATGAACATCGACGAGCATCGCGGTCCCTGGAGTAC-----CCA 296
Db 70 LysCysGlnArgThrIleGlnHisArgHisProPheIleGluTrpAlaArgProPro 89
QY 297 CAGCGC---GGCTGGACATCATCAAGAACTCTCCCGCCTCATCAAGAGCCAC 353
Db 90 GlnProSerGlyValAspLysAlaAsnAlaMetProAlaProArgIleLeuArgThrHis 109
QY 354 CTGCGCTACCGCTTCTCCCTCTGACCTCCACATGGAGACTCCAAGTTCATCATATG 413
Db 110 LeuProThrGlnLeuLeuProProSerPheTrpThrAsnAsnCysLysPheLeuTyrVal 129
QY 414 GCTGCAACCCCAAGGATCTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGGACC 473
Db 130 AlaArgAsnAlaLysAspCysMetValSerTyrThrHisPheTyrArgMetSerGlnVal 149
QY 474 ATGACTACCGAGGACCTTTCAAGAATTCGCCGAGGTTTATGAATGATAAGCTGGGC 533
Db 150 LeuProAspProGlyThrTrpAsnGlnTyrPheGluThrPheIleAsnGlyLysValSer 169
QY 534 TACGCTCTCTGTTGAGACGTCGAGAGTCTCGGAGCACCGCATCGACTCGAACGTG 593
Db 170 TrpGlySerTrpPheAspHisValLysGlyTrpTrpGluIleArgAspArgTyrGlnIle 189
QY 594 CTTTTCCTCAAGATGAGACATGATCGGACCTGGTGACGATGCTGGAGCAGCTGGCC 653
Db 190 LeuPheLeuPheTyrGluAspValLysArgAspProLysArgGluIleGlnLysValMet 209
QY 654 AGATTCTCTGGGGGTCTCC-----TGTGACAAAGCCCGACCTGGAAGCCCTG 698
Db 210 GlnPheMetGlyLysAsnLeuAspGluGluValValAspLysIleValLeuGlnThrSer 229
QY 699 ACGGAGACTGCCACACGCTGTTGACCGCTGCTGCAACGCTGAGCCCTGCCC----- 752
Db 230 PheGluLysMetLysGlu-----AsnProMetThrAsnArgSerThrValProLysSer 247
QY 753 -----GTGGCCGGGGAAGAGTGGGCTGTGGAGGAC 785
Db 248 ValLeuAspGlnSerIleSerProPheMetArgLysGlyThrValGlyAspTrpLysAsn 267
QY 786 ATCTTCACGCTCCCATGAATGAGAGTTTGACTGGTGTATAACAGAAATGGGAAG 845
Db 268 HisPheThrValAlaGlnAsnAspArgPheAspGluIleTyrLysGlnLysMetGlyGly 287
QY 846 TGTGACCTCAGCTT 860
Db 288 ThrSerLeuAsnPhe 292

RESULT 4
S10329
aryl sulfotransferase (EC 2.8.2.1) IV - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S10329; A44932
R:Ozawa, S.; Nagata, K.; Gong, D.; Yamazoe, Y.; Kato, R.
Nucleic Acids Res. 18, 4001, 1990
A:Title: Nucleotide sequence of a full-length cDNA (BST-1) for aryl sulfotransferase from
A:Reference number: S10329; MUID:90326537; PMID:2374726
A:Accession: S10329
A:Molecule type: mRNA
A:Residues: 1-291 <OZA>
A:Cross-references: UNIPROT:P17988; UNIPARC:UPI000000B75; EMBL:X52883; NID:955760; PIDN
R:Verokun, I.; Etheredge, J.L.; Norton, T.R.; Carter, H.A.; Chung, K.H.; Birckbichler, B
Cancer Res. 52, 4779-4786, 1992
A:Title: Characterization of a complementary DNA for rat liver aryl sulfotransferase IV
used hepatocarcinogenesis.
A:Reference number: A44932; MUID:92379794; PMID:1511441
A:Accession: A44932
A>Status: preliminary
A:Molecule type: mRNA; protein
```

A;Residues: 7-291 <YER>  
A;Cross-references: UNIPARC:UPI000017088C; GB:X58640; GB:942994; NID:955764; PIDN:CAA486  
A;Experimental source: liver  
A;Note: sequence extracted from NCBI backbone (NCBIN:112171, NCBIP:112178)  
C;Superfamily: alcohol sulfotransferase  
C;Keywords: sulfotransferase

Alignment Scores:  
Pred. No.: 1,25e-26 Length: 291  
Score: 458.00 Matches: 96  
Percent Similarity: 55.6% Conservative: 57  
Best Local Similarity: 34.9% Mismatches: 104  
Query Match: 10.4% Indels: 18  
Gaps: 4

US-10-768-158-1 (1-2419) x S10329 (1-291)

```
QY 90 GGCCTGGCGTGGCGCCCTCTCGCGGGGAAGATGGAGAGATCGCCAACTTCCCGGTG 149
Db 13 GlyIleProLeuIleLysTyrPheAlaGluThrIleGlyProLeuGlnAsnPheThrAla 32
QY 150 CGGCCCGACGACGTGTGGATCGTCACCTACCCCAAGTCCGCCACAGCTTCTCGAGAG 209
Db 33 TrpProAspAspLeuLeuIleSerThrTyrProLysSerGlyThrThrTrpMetSerGlu 52
QY 210 GTGTCTACTTGTGGAGCCAGCGGCGCTGACCCCGATGAGATCGCTTGTATGAACATGAC 269
Db 53 IleLeuAspMetIleTyrGlnGlyGlyLysLeuGluLysCysGlyArgAlaProIleTyr 72
QY 270 GAGCAGCTCCCGTCTCGAGTACCACACAGCCG-----GGCTCGACATCATC 317
Db 73 AlaArgValProPheLeuGluPheLysCysProGlyValProSerGlyLeuGluThrLeu 92
QY 318 AAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTTCGCCCTCT 377
Db 93 GluGluThrProAlaProArgLeuLeuLysThrHisLeuProLeuSerLeuLeuProGln 112
QY 378 GACCTCCACATGGAGACTCCAAAGTCTATATGGTCCGTCGCAACCCCAAGATCTGGTG 437
Db 113 SerLeuLeuAspGlnLysValLysValIleTyrIleAlaArgAsnAlaLysAspVal 132
QY 438 GTGTCTTATTATCAGTCTCCACCGCTCTCTGGCGACCATGAGCTACCGAGGCACCTTCAA 497
Db 133 ValSerTyrTyrAsnPheTyrAsnMetAlaLysLeuHisProAspProGlyThrTrpAsp 152
QY 498 GAATTCCTCCGAGGTTTATGAATGATAGCTGGGCTACGGCTCTCTGGTTGAGCACGTG 557
Db 153 SerPheLeuGluAsnPheMetAspGlyGluValSerTyrGlySerTrpTyrGlnHisVal 172
QY 558 CAGGAGTCTGGGAGCACCGCATGACTCGAAGCTGCTTTTCTCAAGTATCAAGACATG 617
Db 173 LysGluTrpTrpGluLeuArgHisThrHisProValLeuTyrLeuPheTyrGluAspIle 192
QY 618 CATCGGACCTGGTGCAGCATGGTGGAGCAGCTGGCGAGATTCCTGGGGGTGCTCTGTGAC 677
Db 193 LysGluAsnProLysArgGluIleLysLysIleLeuGluPheLeuGlyArgSerLeuPro 212
QY 678 AAGGCCACAGCTGGAAGCCCTGACGAGCAGCTGC-----CACAGCTGTGTGACCGTGC 731
Db 213 GluGluThrValAspSerIleValHisThrSerPheLysLysMetLysGluAsnCys 232
QY 732 ---TCAAGCTGAGCGCCCTGCC-----GTG 755
Db 233 MetThrAsnTyrThrThrIleProThrGluIleMetAspHisAsnValSerProPheMet 252
QY 756 GCGCGGGGAAGAGTGGGCTGTGGAGGAGGACATCTTCACCGTCTCCATGAATGAGAAGTTT 815
Db 253 ArgLysGlyThrThrGlyAspTrpLysAsnThrPheThrValAlaGlnAsnGluArgPhe 272
QY 816 GACTGTGTGTATAAACAGAAAGTGGAAAGTGTGACCTCAGCTTT 860
Db 273 AspAlaHisTyrAlaLysThrMetThrAspCysAspPheLysPhe 287
```



```

157 SerPheLeuGluLysPheMetAlaGlyGluValSerTyrGlySerTrpTyrGlnHisVal 176
558 CAGGAGTTCTGGGAGCACCAGTGCACGCTGCAAGCTGCTTTTCTCAAGTATGAAGACATG 617
177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
618 CATCGGACGCTGGTGACGATGGTGGAGCAGCTGGCCAGATTCCTGGGGGTGTCCTGTGAC 677
197 LysGluAenProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
678 AAGGCCACGCTGGAAGCCCTGACGGAGCACTGC----- 710
217 GluGluThrValAspLeuMetValGluHisThrSerPheLysGluMetLysLysThrPro 236
711 -----CACACGCTGGTGACCACTGTGCTGCAACGCTGAGGCC 746
237 MetThrAsnTyrThrThrValArgGluPheMetAspHisSerIleSer----- 253
747 CTGCCC---GTGGCGCCGGGAGAGATTGGCTGTGGAGGAGCATCTTCACGCTGCCATG 803
254 --ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGln 272
804 AATCAGAACTTGACTTGGTGTATATAACAGAGATGGGGAAGTCGACCTCACGTTT 860
273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 7
G01843
aryl sulfotransferase - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
C:Accession: G01843
R:Zhu, X.
submitted to the EMBL Data Library, May 1995
A:Reference number: G08606
A:Accession: G01843
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: mRNA
A:Residues: 1-295 <ZHU>
A:Cross-references: UNIPROT:P50226; UNIPARC:UPI0000154B6D; EMBL:U28169; NID:g8198
C:Superfamily: alcohol sulfotransferase

Alignment Scores:
Pred. No.: 1,78e-26 Length: 295
Score: 456.00 Matches: 94
Percent Similarity: 54.5% Conservative: 58
Best Local Similarity: 33.7% Mismatches: 101
Query Match: 10.4% Indels: 10, 26
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x G01843 (1-295)

```



QY	270	GAGCAGCTCCGGTCTCGAGTACCCACAGCCG-----GGCTGGACATCATC	317
Db	77	MetArgValProPheLeuGluPheLysAlaProGlyIleProSerGlyMetGluThrLeu	96
QY	318	AAGGAAGTACGCTCTCCCGCTCATCAAGAGCCACTGCCCTACCGCTTCTGCGCTCT	377
Db	97	LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuProGln	116
QY	378	GACCTCCACAATGGAGACTCCAAAGTCACTATATATGGTCGCAACCCCAAGATCTGGG	437
Db	117	ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspValAla	136
QY	438	GTGCTTATTATCAGTTCACCGCTCTCTCGGACCATGAGTACCGAGCCACCTTCAA	497
Db	137	ValSerTyrTyrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp	156
QY	498	GAATTTCTGCGGAGTTTATGAATGATGAAGCTGGCTACGGCTCCCTGGTTTGAGCAGTG	557
Db	157	SerPheLeuGluLysPheMetValGlyGluValSerTyrGlySerTrpTyrGlnHisVal	176
QY	558	CAGGAGTTCTGGAGCACCGCATGGACTCGAAGTCTTTTCTCAAGTATGAAGACATG	617
Db	177	GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet	196
QY	618	CATCGGGACCTGGTACGATGTGTGAGCAGCTGGCCAGATTCTGGGGGTGCTGTGAC	677
Db	197	LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro	216
QY	678	AAGGCCAGCTGGAAGCCTGACGGACACTGC-----CACCAGCTGGTGACCAG	728
Db	217	GluGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysAsnPro	236
QY	729	TGCTGCAACGCTGAGCGCCCTGCC-----GTG	755
Db	237	MetThrAsnTyrThrThrValProGlnGluPheMetAspHisSerIleSerProPheMet	256
QY	756	GCCTGGGAGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCCATGATGAGAAGTTT	815
Db	257	ArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPhe	276
QY	816	GACTTGGTGTATAACAGAAAGTGGAAAGTGTGACCTCACGTTT	860
Db	277	AspAlaAspTyrAlaLysLysMetAlaGlyCysSerLeuThrPhe	291
RESULT 9			
S52794			
aryl sulfotransferase (EC 2.8.2.1) - human			
C:Species: Homo sapiens (man)			
C>Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004			
C:Accession: S52794			
R:Ozawa, S.; Nagata, K.; Shimada, M.; Ueda, M.; Tsuzuki, T.; Yamazoe, Y.; Kato, R.			
submitted to the EMBL Data Library, March 1994			
A:Description: Primary structures and properties of two related aryl sulfotransferases			
A:Reference number: S52791			
A:Accession: S52794			
A>Status: preliminary			
A:Molecule type: mRNA			
A:Residues: 1-295 <OZA>			
A:Cross-references: UNIPROT:P50225; UNIPARC:UPI0000136228; EMBL:X78283; NID:g758596; PID			
C:Superfamily: alcohol sulfotransferase			
C:Keywords: sulfotransferase			
Alignment Scores:			
Pred. No.:	4-27e-26	Length:	295
Score:	451.00	Matches:	93
Percent Similarity:	54.5%	Conservative:	57
Best Local Similarity:	33.8%	Mismatches:	107
Query Match:	10.2%	Indels:	18
DB:	2	Gaps:	3

US-10-768-158-1 (1-2419) x S52794 (1-295)

QY	90	GCGTGTGCGCTGCGCCCTTCTGCGCGGAAGATGAGAGATCCCAACTTCCCGGTG	149
Db	17	GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla	36
QY	150	CGGCCCAGCAGCTGTGGGATCGTACCTACCCCAAGTCCGCGCACAGCTGTGTGACGAG	209
Db	37	ArgProAspAspLeuLeuIleSerThrTyrProLysSerGlyThrThrTrpValSerGln	56
QY	210	GTGGTCTACTGTGAGCCAGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC	269
Db	57	IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysHisArgAlaProIlePhe	76
QY	270	GAGCAGCTCCGGTCTCGAGTACCCACAGCCG-----GGCTGGACATCATC	317
Db	77	MetArgValProPheLeuGluPheLysAlaProGlyIleProSerGlyMetGluThrLeu	96
QY	318	AAGGAAGTACGCTCTCCCGCTCATCAAGAGCCACTGCCCTACCGCTTCTGCGCTCT	377
Db	97	LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuProGln	116
QY	378	GACCTCCACAATGGAGACTCCAAAGTCACTATATATGGTCGCAACCCCAAGATCTGGG	437
Db	117	ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspValAla	136
QY	438	GTGCTTATTATCAGTTCACCGCTCTCTCGGACCATGAGTACCGAGCCACCTTCAA	497
Db	137	ValSerTyrTyrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp	156
QY	498	GAATTTCTGCGGAGTTTATGAATGATGAAGCTGGCTACGGCTCCCTGGTTTGAGCAGTG	557
Db	157	SerPheLeuGluLysPheMetValGlyGluValSerTyrGlySerTrpTyrGlnHisVal	176
QY	558	CAGGAGTTCTGGAGCACCGCATGGACTCGAAGTCTTTTCTCAAGTATGAAGACATG	617
Db	177	GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet	196
QY	618	CATCGGGACCTGGTACGATGTGTGAGCAGCTGGCCAGATTCTGGGGGTGCTGTGAC	677
Db	197	LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro	216
QY	678	AAGGCCAGCTGGAAGCCTGACGGACACTGC-----CACCAGCTGGTGACCAG	728
Db	217	GluGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysAsnPro	236
QY	729	TGCTGCAACGCTGAGCGCCCTGCC-----GTG	755
Db	237	MetThrAsnTyrThrThrValProGlnGluPheMetAspHisSerIleSerProPheMet	256
QY	756	GCCTGGGAGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCCATGATGAGAAGTTT	815
Db	257	ArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluArgPhe	276
QY	816	GACTTGGTGTATAACAGAAAGTGGAAAGTGTGACCTCACGTTT	860
Db	277	AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuThrPhe	291
RESULT 10			
JC5248			
aryl sulfotransferase (EC 2.8.2.1) HAST2 - human			
N:Alternate names: aryl sulfotransferase 1; phenol sulfotransferase; sulfokinase			
C:Species: Homo sapiens (man)			
C>Date: 13-Mar-1997 #sequence_revision 25-Aug-1997 #text_change 09-Jul-2004			
C:Accession: JC5248; JN0714; JN0714; JN0714; JN0714; JN0714			
R:Dooley, T.P.; Huang, Z.			
Biochem. Biophys. Res. Commun. 228, 134-140, 1996			
A:Title: Genomic organization and DNA sequences of two human phenol sulfotransferase genes			
A:Reference number: JC5248; MUID:97069665; PMID:8912648			
A:Accession: JC5248			
A>Status: nucleic acid sequence not shown			
A:Molecule type: mRNA			
A:Residues: 1-295 <DOO>			
A:Cross-references: UNIPROT:P50225; UNIPARC:UPI000000122E; GB:U71086; NID:g1575787; PIDN			
R:Zhu, X.; Veronese, M.E.; Bernard, C.C.A.; Sansom, L.N.; McManus, M.E.			

R;Hwang, S.R.; Kohn, A.B.; Hook, V.Y.H.  
 Biochem. Biophys. Res. Commun. 207, 701-707, 1995  
 A;Title: Molecular cloning of an isoform of phenol sulfotransferase from human  
 A;Reference number: JC2523; MUID:95169114; PMID:7864863  
 A;Accession: JC2523  
 A;Molecule type: mRNA  
 A;Residues: 1-295 <HWA>  
 A;Cross-references: UNIPROT:P50225; UNIPARC:UPI000016A1D1; GB:U26309; NID:G9847  
 A;Experimental source: hippocampus

```

CysSuperfamily: glycine sulfoxylase
CysKeywords: sulfoxyltransferase
F:70/Active site: Cys #status Predicted

Alignment Scores:
Pred. No.:          6, 06e-26          Length:          295
Score:             449.00             Matches:         93
Percent Similarity: 54.2%             Conservative:    56
Best Local Similarity: 33.8%          Mismatches:     108
Query Match:       10.2%             Gaps:           18
DB:                2                Gaps:           3

US-10-768-158-1 (1-2419) x JC2523 (1-295)

QY 90 GCGGTGGCGGTGCCGCCCTTGTGCCGCGGAAGATGGAGGAGATGCCCAACTTCGGGTG 149
||||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36

QY 150 CGGCCCGAGCGGTGTGGATGCTCACTACCCCAAGTCGCGCACACAGCTGTGCTCAGGAG 209
||||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 37 ArgProAspAspLeuIleSerThrTyrProLysSerGlyThrThrProValSerGln 56

QY 210 GTGGTCTACTGGTGAGCGAGCGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysHisArgAlaProIlePhe 76

QY 270 GAGCAGCTCCGGTCTCTGGAGTACCCACAGCCG-----GGCCTGGACATCATC 317
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 27 GAGCAGCTCCGGTCTCTGGAGTACCCACAGCCG-----GGCCTGGACATCATC 317
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :

```



A: Experimental source: liver  
C: Comment: This enzyme mediates the transfer of SO3- from 3'-phosphoadenosine 5'-phosphocysteine to macromolecules.  
C: Superfamily: alcohol sulphotransferase  
C: Keywords: sulphotransferase

Alignment Scores:  
Pred. No.: 9 41e-26 Length: 299  
Score: 446.50 Matches: 95  
Percent Similarity: 55.7% Conservative: 52  
Best Local Similarity: 36.0% Mismatches: 98  
Query Match: 10.1% Indels: 19  
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x J05884 (1-299)

```
QY 126 GAGGAGATCGCAACTTCCCGTGGCGCCACGACGAGTGTGGATCGTCACCTACCCCAAG 185
DB 29 GluLysIleGluGluPheGlnSerArgProCysAspIleValIleProThrTyrProLys 48
QY 186 TCCGGACACAGTTCGTGACGAGGTGTCTACTTGTGTGACGCGCGCTGACCCCGAT 245
DB 49 SerGlyThrTrpLeuSerGluIleValAspMetValLeuAsnAspGlyAsnValGlu 68
QY 246 GAGATCGGCTTGATGAACATGACGACGAGTCCCGTCTCGAGTACCCACAGCG--- 302
DB 69 LysCysLysArgAspValIleThrSerLysValProMetLeuGluGlnAsnValProGly 88
QY 303 -----GGCTGGACATCATCAAGAACTGACCTCCCGCTCATCAAGAGC 350
DB 89 AlaArgSerGlyValGluLeuLeuLysLysThrProSerProArgIleIleLysThr 108
QY 351 GACCTGCCCTACCGCTTCTGCGCTGACCTCCACATGGAGACTCCAGGTCTACTAT 410
DB 109 HisLeuProIleAspLeuLeuProLysSerPheThrAspAsnLysCysLysMetIleTyr 128
QY 411 ATGGCTCGCAACCCCAAGATCTGGTGTCTTATTATCATCTCCACCGCTCTCGCG 470
DB 129 LeuAlaArgAsnGlyLysAspValAlaValSerTyrTyrHisPheAspLeuMetAsnLeu 148
QY 471 ACCATGAGTACCGGACCTTTCAAGAAATCTCGCGAGGTATTGAATGAAGCTG 530
DB 149 IleGlnProLeuProGlyThrTrpGluGluLysPheLeuAlaGlyAsnVal 168
QY 531 GCTACGGCTCTGTTTACGACGTCGAGGATCTCGGAGCAGCGCTGACCTCGAAC 590
DB 169 AlaTyrGlySerTrpPheAspHisValLysSerTrpTrpGluLysArgGluGlyHisPro 188
QY 591 GTGCTTTTCTCAAGTATGAACATGTCATCGGACCTGGTACGATGGTGAGCAGCTG 650
DB 189 IleLeuPheLeuTyrTyrGluAspLeuLysLysAsnProLysLysGluIleLysIle 208
QY 651 GCCAGATTCCTGGGGTGTCTGTGACAGGCCAGCTCGAAGCCCTGACGAGCACTGC 710
DB 209 AlaAsnPheLeuAspLysThrLeuAspGluHisThrLeuGluArgIleValHisThr 228
QY 711 CACCAGCTGGT-----GACCAGTGTGCAACGCTGAGGCCCTGCC----- 752
DB 229 SerPheGluValMetLysAspAsnProLeuValAsnTyrThrHisLeuProThrGluIle 248
QY 753 -----GTGGCGGGGGAAGATTGGGCTGTGGAGGACATC 788
DB 249 MetAspHisSerLysSerProPheMetArgLysGlyValValGlyAspTrpLysAsnTyr 268
QY 789 TTCACGCTCCATGAATGAGAGTTTGATCTGGTGTATATAACAGAAAGTGAAGTGT 848
DB 269 PheThrMetThrGlnSerGluLysPheAspAlaIleTyrLysLysLysLeuSerGlyThr 288
QY 849 GACCTCAGCTTT 860
DB 289 ThrLeuGluPhe 292
```

RESULT 14  
G02924

phenol sulphotransferase subunit - crab-eating macaque (fragment)  
C: Species: Macaca fascicularis (crab-eating macaque)  
C: Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 09-Jul-2004  
C: Accession: G02924

R: Ogura, K.  
submitted to GenBank, May 1996  
A: Reference number: H01975  
A: Accession: G02924  
A: Status: preliminary; translated from GB/EMBL/DBJ  
A: Molecule type: mRNA  
A: Residues: 1-295 <OGU>  
A: Cross-references: UNIPROT:P52846; UNIPARC:UPI000013622B; GB:D85514; NID:g1339922; PIDN  
C: Superfamily: alcohol sulphotransferase

Alignment Scores:  
Pred. No.: 1 02e-25 Length: 295  
Score: 446.00 Matches: 93  
Percent Similarity: 54.2% Conservative: 56  
Best Local Similarity: 33.8% Mismatches: 108  
Query Match: 10.1% Indels: 18  
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x G02924 (1-295)

```
QY 90 GCGTGGCGCTGCGCCCTTCTCGCGGGAAGATGGAGGATCGCCAACTTCCCGGTG 149
DB 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheArgAla 36
QY 150 CGGCCACGACGAGTGTGGATCGTCACCTACCCCAAGTCCGACAGCTTCTGTCAGAG 209
DB 37 ArgProAspAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpValSerGln 56
QY 210 GTGTCTACTTGGTGGACGCGGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
DB 57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysArgAlaProIlePhe 76
QY 270 GAGCAGCTCCCGTCTCGAGTACCACAGCCG-----GGCTGGACATCATC 317
DB 77 MetArgValProPheLeuGluPheLysValProGlyIleProSerGlyMetGluThrLeu 96
QY 318 AAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCTACCGCTTCTGCGCTCT 377
DB 97 LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuProGln 116
QY 378 GACCTCCACATGGAGACTCCAAAGTCTATATGGCTCGCAACCCCAAGATCTGTG 437
DB 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspValAla 136
QY 438 GTGTCTTATTACGTTCCACCGCTCTCTCGCGACCATGAGCTACCGAGGACCTTCAA 497
DB 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp 156
QY 498 GAATTCCTCCGAGGTTTATGAATGATAGCTGGGCTACGGCTCTGTTGTTGAGCAGTG 557
DB 157 SerPheLeuGluLysPheMetAlaGlyGluValSerTyrGlySerTyrTrpGlnHisVal 176
QY 558 CAGGAGTCTCGGACACCGCATGACCTGCAACGCTTTTCTCAAGTATGAACACATG 617
DB 177 GlnGluTrpTrpGluLeuSerHisThrHisProValLeuTyrLeuPheTyrGluAspMet 196
QY 618 CATCGGACCTGCTGACGATGGTGAGCAGCTGGCGAGATTCTCGGGGTGTCTGTGAC 677
DB 197 LysGluAsnProLysArgGluIleTrpLysIleLeuGluPheValGlyArgSerLeuPro 216
QY 678 AAGGCCAGCTGGAAGCCCTGACGAGCATCTGC-----CACCAGCTGTGGACCCAG 728
DB 217 GluGluThrValAspLeuMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236
QY 729 TCGTCGAACGCTGAGGCCCTGCC-----GTG 755
DB 237 MetAlaAsnTyrThrIleProGlnGluLeuMetAspHisSerIleSerProPheMet 256
QY 756 GCGCGGGAAGAGTTGGGCTGTGGAAGGACATCTTTCACCGTCTCCATGAATGAGAAGTTT 815
```

Db 257 ArgLysGlyMetThrGlyAspTrpLysThrThrPheThrValAlaGlnAsnGluHisPhe 276  
QY 816 GACTTGGTGTATAAACAGAGATGGGAAAGTGTGACCTCACGTTT 860  
Db 277 AspValAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291  
RESULT 15  
S29045  
estron sulfotransferase (EC 2.8.2.4) - bovine  
N:Alternate names: estrogen sulfotransferase  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 25-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S29045; S18222  
R:Nash, A.R.; Glenn, W.K.; Moore, S.S.; Kerr, J.; Thompson, A.R.; Thompson, E.O.P.  
Aust. J. Biol. Sci. 41, 507-516, 1988  
A:Title: Oestrogen sulfotransferase: molecular cloning and sequencing of cDNA for the bovine  
A:Reference number: S29045; MUID: 90147277; PMID: 3271383  
A:Accession: S29045  
A:Molecule type: mRNA  
A:Residues: 1-295 <NAS>  
A:Cross-references: UNIPROT:P19217; UNIPARC:UPI0000136223; EMBL:M54942; NID:g163462; PID  
A:Experimental source: placenta  
R:Adams, J.B.  
Biochim. Biophys. Acta 1076, 282-288, 1991  
A:Title: Enzymic synthesis of steroid sulphates XVII. On the structure of bovine estrogen  
A:Reference number: S13486; MUID: 91152101; PMID: 1300200  
A:Accession: S18222  
A:Molecule type: protein  
A:Residues: 146-160, 206-220, 257-259 <ADA>  
A:Cross-references: UNIPARC:UPI0000175826; UNIPARC:UPI0000175827; UNIPARC:UPI0000175828  
A:Experimental source: placenta  
C:Complex: homodimer  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: homodimer; phosphoprotein; sulfotransferase  
Alignment Scores:  
Pred. No.: 1,02e-25 Length: 295  
Score: 446.00 Matches: 100  
Percent Similarity: 53.8% Conservative: 57  
Best Local Similarity: 34.2% Mismatches: 111  
Query Match: 10.1% Indels: 24  
DB: 2 Gaps: 5  
US-10-768-158-1 (1-2419) x S29045 (1-295)  
QY 48 AGCACCCCGGGGAGTTCGAGACCAAGTACTTCGAGTTCCTCGAGTTCCTCGGCTGGCGGCGGCC 107  
Db 3 SerSerLysProSerPheSerAspTyrPheGlyLysLeuGlyGlyle-----Pro 19  
QY 108 TTCCTGCGCGGGGAAGATGAGAGAGATCGCCAAAC-----TTCCTGCGGTGCGGCCCGCAG 158  
Db 20 MetTyrLysLysPheIleGluGlnPheHisAsnValGluGluPheGluAlaArgProAsp 39  
QY 159 GAGCTGTGATCTCATCCCAAGTCCGACCAAGTTCGCTGCTGCGAGGAGTGGTCTAC 218  
Db 40 AspLeuValIleValThrTyrProLysSerGlyThrThrTripleSerGluIleIleCys 59  
QY 219 TTGCTGAGCCAGCGGCGCTGACCCCGATGATGATCGGCTTGTATGACATCCAGCAGCAGCTC 278  
Db 60 MetIleTyrAsnAsnGlyAspValGluLysCysLysGluAspValIlePheAsnArgVal 79  
QY 279 CCGGTCTCTG-----GAGTACCCACACCGCGGCTGGACATCATCAAGGAACATG 326  
Db 80 ProTyrLeuGluCysSerThrGluHisValMetLysGlyValLysGlnLeuAsnGluMet 99  
QY 327 ACCTCTCCCGCTCATCAAGACCCACCTGCCCTACCGCTTCTGCGCTCTGACCTCCAC 386  
Db 100 AlaSerProArgIleValLysSerHisLeuProValLysLeuLeuProValSerPheTyr 119  
QY 387 AATGGAGACTCCAAAGTTCATATATATGCTCGCAACCCCAAGATCTGTGTGTCTTAT 446  
Db 120 GluLysAsnCysLysIleIleTyrLeuSerArgAsnAlaLysAspValValSerTyr 139

QY 447 TATCAGTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGGCACCTTTCAAGAAATTCG 506  
Db 140 TyrPheLeuIleLeuMetValThrAlaIleProAspProAspSerPheGlnAspPheVal 159  
QY 507 CGGAGGTTTATGAATGAATAGCTGGCTACGGCTCCTGTGTTTGGACACGTCGAGGAGTTC 566  
Db 160 GluLysPheMetCaspGlyGluValProTyrGlySerTrpPheGluHisThrLysSerTrp 179  
QY 567 TGGGAGCACCGCATGGACTCGAACGTCGTTTTCTCAAGTATGAAGACATGCATCGGAC 626  
Db 180 TrpGluLysSerLysAsnProGlnValLeuPheLeuPheTyrGluAspMetLysGluAsn 199  
QY 627 CTGGTGACAGTGTGAGCAGCTGCCAGATTCCTGGGGTGTCTCTGTGACAGGCCGAC 686  
Db 200 IleArgLysGluValMetLysLeuLeuGluPheLeuGlyArgLysAlaSerAspGluLeu 219  
QY 687 CTGGAAGCCCTGACCGACACTGCCAC----- 713  
Db 220 ValAspLysIleIleLysHisThrSerPheGlnGluMetLysAsnAsnProSerThrAsn 239  
QY 714 -----CAGCTGCTGGACAGCTGCTGCAACGCTGAGGCCCTGCC-----GTGGCGCGGGA 764  
Db 240 TyrThrThrLeuProAspGluValMetAsnGlnLysValSerProPheMetArgLysGly 259  
QY 765 AGAGTTGGCTGTGAAGACATCTCACCGTCTCCATGAATGAGAAGTTTGAATGTTGG 824  
Db 260 AspValGlyAspTyrLysAsnHisPheThrValAlaLeuAsnGluLysPheAspMetHis 279  
QY 825 TATAACAGACAGATGGGAAGTGTGACCTCACGTTT 860  
Db 280 TyrGluGlnGlnMetLysGlySerThrLeuLysPhe 291  
RESULT 16  
S52399  
aryl sulfotransferase (EC 2.8.2.1) - human  
C:Species: Homo sapiens (man)  
C:Date: 08-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 18-Jun-1999  
C:Accession: S52399  
R:Jones, A.L.; Hagen, M.; Coughtrie, M.W.H.; Roberts, R.C.; Glatt, H.R.  
submitted to the EMBL Data Library, February 1995  
A:Description: Human platelet phenolsulfotransferases: cDNA cloning, stable expression  
A:Reference number: S52398  
A:Accession: S52399  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-295 <JON>  
A:Cross-references: UNIPARC:UPI000016B4D7; EMBL:X84654; NID:g671641; PIDN:CAA59147.1; PI  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: sulfotransferase  
Alignment Scores:  
Pred. No.: 1,45e-25 Length: 295  
Score: 444.00 Matches: 92  
Percent Similarity: 54.2% Conservative: 57  
Best Local Similarity: 33.5% Mismatches: 108  
Query Match: 10.1% Indels: 18  
DB: 2 Gaps: 3  
US-10-768-158-1 (1-2419) x S52399 (1-295)  
QY 90 GCGTGGCGGCTGCCCGCCCTTCTGCGCGGGAAGATGGAGGATCGCCAACTTCCGGTG 149  
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36  
QY 150 CGGCCCGGAGCTGTGGATGTCACCTACCCCAAGTCCGACCCAGCTTGTGTGAGGAG 209  
Db 37 ArgProAspAspLeuIleSerThrTyrProLysSerGlyThrThrTrpValSerGln 56  
QY 210 GTGGTCTACTTGGTGGCCAGCGGCTGACCCCGATGATGATCGGCTTGTATGAACATCAG 269  
Db 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysHisArgAlaProIlePhe 76



Alignment Scores:				
Pred. No.:	7,01e-25	Length:	295	
Score:	435.00	Matches:	93	
Percent Similarity:	53.1%	Conservative:	62	
Best Local Similarity:	31.8%	Mismatches:	105	
Query Match:	9.9%	Indels:	32	
DB:	2	Gaps:	4	
US-10-768-158-1 (1-2419) x I73679 (1-295)				
QY	60	GAGTTTCGACGACGAAGTACTTCGAGTTCATGGCGTGGCGCTGCGCGCTTCTGCGCGGG	119	
Db	7	GlutyrTyrrGluValPheGlyAspPheHisGlyValLeuValAspLysLeuPheThrLys	26	
QY	120	AAGATGGAGGAGATGCCAACTTCGCGGTGCGCGCCAGCGACGTGTGGATCGTCACTAC	179	
Db	27	TyrTrpGluAspValGluThrPheSerAlaAargProAspAspLeuValValThrTyr	46	
QY	180	CCCAAGTCCGGCACCAGCTTGCTGCGAGGAGTGGTCTACTTGGTGAGCCAGGCGGTGAC	239	
Db	47	ProLysSerGlySerThrTrpIleGlyGluIleValAspMetIleTyrLysGluGlyAsp	66	
QY	240	CCCGATGAGATCGGCTTGATGAACATCGACGACGTCCCGTCTCTGGAGTACCCACAG	299	
Db	67	ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn	86	
QY	300	CCG-----GGCTGGACATCATCAAGGAACCTGACCTCTCCCGCTCATCAAG	347	
Db	87	GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleValLys	106	
QY	348	AGCCACCTGCGCTACCGCTTCTGCGCTCTGACCTCCACAAATGGAGACTCCAAGTCAATC	407	
Db	107	ThrHisLeuProAlaLysLeuLeuProAlaSerPheTrpGluLysAsnCysLysIleIle	126	
QY	408	TATATGGCTCGCAACCCCAAGGATCTGGTGTGTTATTATTCAGTTCCACCGCTCTCG	467	
Db	127	TyrLeuCysArgAsnAlaLysAspValValSerTyrTyrPhePheLeuIleIle	146	
QY	468	CGACCATGAGCTACCGAGCACCTTTCAAGAATCTCGCGGAGGTTTATGAATGAAG	527	
Db	147	LysSerTyrrProAsnProLysSerPheSerGluPheValGluLysPheMetGluGln	166	
QY	528	CTGGCTACGGCTCTCTGTTTACGACGTGACGAGGTCTGGAGCACCGCATGACTCG	587	
Db	167	ValProTyrGlySerTrpTyrAspHisValLysSerTrpTrpGluLysSerLysAsnSer	186	
QY	588	AACGTGCTTTTCTCAAGTATCAAGACATGCATCGGACCTGCTGACGATGTGGAGCAG	647	
Db	187	ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgGluValValLys	206	
QY	648	CTGGCCAGATTCCTGGGGGTGCTCTGTGACAAAGGCCAGCTCGAAGCCCTGACGGAGCAC	707	
Db	207	LeuIleGluPheLeu-----GluArgAspProLeuAlaGluLeu	219	
QY	708	TCCACACAGCTGGT-----GACCAAGTGGTGCAC	737	
Db	220	ValAspLysIleIleGlnHisSerPheGlnGluMetLysAsnAsnProCysThrAsn	239	
QY	738	GCTAGGCCCTCCCT-----GTGGCCGGGGA	764	
Db	240	TyrSerMetLeuProGluThrMetIleAspLeuLysValSerProPheMetArgLysGly	259	
QY	765	AGATGTGGCTGTGGAAGACATCTTACCGTCTCCATGAATGAGAAGTTTCACTTGGTG	824	
Db	260	IleValGlyAspTrpArgAsnHisPheProGluAlaLeuArgGluArgPheGluHis	279	
QY	825	TATAACAGACAGATGGAAAGTGTGACCTCAGCTTT	860	
Db	280	TyrGlnArgHisMetLysAspCysProValThrPhe	291	
RESULT 19				
JC2229				
estrogen sulfotransferase (EC 2.8.2.-) - human				

C:Species: Homo sapiens (man)				
C>Date: 28-Aug-1985 #sequence revision 07-Oct-1994 #text_change 09-Jul-2004				
C:Accession: JC2229; I38249; I56607				
R:Aksoy, I.A.; Wood, T.C.; Weinshilboum, R.				
Biochem. Biophys. Res. Commun. 200, 1621-1629, 1994				
A>Title: Human liver estrogen sulfotransferase: identification by cDNA cloning and expres				
A:Reference number: JC2229; MUID:94242031; PMID:8185618				
A:Accession: JC2229				
A:Molecule type: mRNA				
A:Residues: 1-294 <AKS>				
A:Cross-references: UNIPROT:P49888; UNIPARC:UPI00000623C8; GB:U08098; NID:g488282; PIDN:				
A:Experimental source: liver				
R:Her, C.; Aksoy, I.A.; Kimura, S.; Brandriff, B.F.; Wasmuth, J.J.; Weinshilboum, R.M.				
Genomics 29, 16-23, 1995				
A>Title: Human estrogen sulfotransferase gene (STE): Cloning, structure and chromosomal				
A:Reference number: I38249; MUID:96079087; PMID:8530066				
A:Accession: I38249				
A>Status: preliminary				
A:Molecule type: DNA				
A:Residues: 1-294 <RES>				
A:Cross-references: UNIPARC:UPI00000623C8; EMBL:U20521; NID:g1079511; PIDN:AAC50286.1; PID:				
R:Palany, C.N.; Krasnykh, V.; Falany, J.L.				
J. Steroid Biochem. Mol. Biol. 52, 529-539, 1995				
A>Title: Bacterial expression and characterization of a cDNA for human liver estrogen s				
A:Reference number: I56607; MUID:95298714; PMID:7779757				
A:Accession: I56607				
A>Status: preliminary; translated from GB/EMBL/DDBJ				
A:Molecule type: mRNA				
A:Residues: 1-294 <RES>				
A:Cross-references: UNIPARC:UPI00000623C8; GB:S77383; NID:g998887; PIDN:AAB34601.1; PID:				
C:Comment: This enzyme catalyzes the sulfate conjugation of estrogen.				
C:Genetics:				
A:Gene: GDB:STE				
A:Cross-references: GDB:342104; OMIM:600043				
A:Map position: 16pter-16qter				
A:Introns: 49/1, 91/1, 123/3, 166/1, 197/3; 258/1				
C:Superfamily: alcohol sulfotransferase				
C:Keywords: sulfotransferase				
Alignment Scores:				
Pred. No.:	8,35e-25	Length:	294	
Score:	434.00	Matches:	95	
Percent Similarity:	53.0%	Conservative:	56	
Best Local Similarity:	33.3%	Mismatches:	116	
Query Match:	9.9%	Indels:	18	
DB:	2	Gaps:	3	
US-10-768-158-1 (1-2419) x JC2229 (1-294)				
QY	60	GAGTTTCGACGACGAAGTACTTCGAGTTCATGGCGTGGCGCTGCGCGCTTCTGCGCGGG	119	
Db	6	AspTyrrGluLysPheGluGluValHisGlyIleLeuMetTyrLysAspPheValLys	25	
QY	120	AAGATGGAGGAGATCGCCAACTTCGCGGTGCGCGCCAGCGAGCTGTGGATCGTCACTAC	179	
Db	26	TyrTrpAspAsnValGluAlaPheGlnAlaArgProAspAspLeuAlaIleThrTyr	45	
QY	180	CCCAAGTCCGGCACCAGCTTCTGCGAGGAGTGTCTACTTGGTCAGCCAGCGCGCTGAC	239	
Db	46	ProLysSerGlyThrTrpValSerGluIleValTyrMetIleTyrLysGluGlyAsp	65	
QY	240	CCCGATGAGATCGGCTTGATGAACATCGACGACGAGCTCCCGGTCTCGAGTACCACAG	299	
Db	66	ValGluLysCysLysGluAspValIlePheAsnArgIleProPheLeuGluCysArgLys	85	
QY	300	CCG-----GGCTGGACATCATCAAGGAACCTGACCTCTCCCGCTCGAGTACCACAG	347	
Db	86	GluAsnLeuMetAsnGlyValLysGlnLeuAspGluMetAsnSerProArgIleValLys	105	
QY	348	AGCCACCTGCGCTACCGCTTCTGCGCTCTGACCTCCACAAATGGAGACTCCAAGTGTATC	407	
Db	106	ThrHisLeuProGluLeuProAlaSerPheTrpGluLysAspCysLysIleIle	125	











A:Molecule type: protein  
A:Residues: 23-36;38-47;79-96;134-144;207-226;234-248 <SCW>  
A:CROSS-references: UNIPARC:UPI0000175829; UNIPARC:UPI000017582A; UNIPARC:UPI000017582B;  
C:Comment: This enzyme belongs to a family of cytosolic enzymes that catalyze the transfe  
C:Genetics:  
A:Gene: psc  
A:Introns: 50/1; 93/3; 124/3; 167/1; 198/3; 259/1  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: sulfotransferase

Alignment Scores:  
Pred. No.: 3 91e-23 Length: 294  
Score: 412.00 Matches: 91  
Percent Similarity: 53.5% Conservative: 56  
Best Local Similarity: 33.1% Mismatches: 108  
Query Match: 9.4% Indels: 20  
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x JCS000 (1-294)

```
QY 69 AGCAAGTACTTCGAGTTCACGTCGGCTGCCCGCTTCTGCCGCGGGAAGATGGAG 128
   ::::: ::::: ::::: ::::: ::::: :::::
Db 12 AlaLeuTyrrVallys-----GlylleProLeulleLysTyrrPheAlaGluAlaLeuGly 29
   ::::: ::::: ::::: ::::: ::::: :::::
QY 129 GAGATGCCCAACTTCCCGTGGCGGCCAGGAGCTGTGGATCGTACCTACCCCAAGTCC 188
   ::::: ::::: ::::: ::::: ::::: :::::
Db 30 ProLeuGluSerPheGluAlaTrpProAspAspLeulleSerThrTyrrProLysSer 49
   ::::: ::::: ::::: ::::: ::::: :::::
QY 189 GGCACCAAGTCTGCTGAGGAGTGTCTACTTGTGGAGCGCGCTCACCCCGATGAG 248
   ::::: ::::: ::::: ::::: ::::: :::::
Db 50 GlyThrThrTrpValSerGluLeuLeuAspLeulleTyrrGlnGluGlyAspLeuGluLys 69
   ::::: ::::: ::::: ::::: ::::: :::::
QY 249 ATCGGTTGATGAACATCAGCAGCAGCTCCGGTCTCGAGTACCACAGCCG----- 302
   ::::: ::::: ::::: ::::: ::::: :::::
Db 70 CysGlnArgAlaProValPheLeuArgValProPheLeuGluPheSerAlaProGlyVal 89
   ::::: ::::: ::::: ::::: ::::: :::::
QY 303 -----GGCTGGACATCATCAAGAACTGACTCTCCCGCTCATCAAGACCACTG 356
   ::::: ::::: ::::: ::::: ::::: :::::
Db 90 ProThrGlyValGluLeuLeuLysAspThrProAlaProArgLeuLeuLysThrHisLeu 109
   ::::: ::::: ::::: ::::: ::::: :::::
QY 357 CCTACCGCTTCTGCGCTCTGACCTCCCAATGGAGACTCCAAAGTCTATATATGGCT 416
   ::::: ::::: ::::: ::::: ::::: :::::
Db 110 ProLeuAlaLeuLeuProLysThrLeuLeuAspGlnLysVallyleTyrrleAla 129
   ::::: ::::: ::::: ::::: ::::: :::::
QY 417 CGCAACCCCAAGATCTGTGTGTCTTATATCATGTTCCACGCTCTCTCGGACCATG 476
   ::::: ::::: ::::: ::::: ::::: :::::
Db 130 ArgAsnAlaLysAspValAlaValSerTyrrHisPheTyrrArgMetAlaLysValHis 149
   ::::: ::::: ::::: ::::: ::::: :::::
QY 477 AGCTACCGAGGACCTTCAAGAACTTCCGCGAGGTTTATGAATGATGAGCTGGGCTAC 536
   ::::: ::::: ::::: ::::: ::::: :::::
Db 150 ProAspProGlyThrTrpAspSerPheLeuGluLysPheMetAlaGlyGluValCysTyrr 169
   ::::: ::::: ::::: ::::: ::::: :::::
QY 537 GGCTCTGTTTGGACACGTCGAGGAGTTCGGGACCGCGCATGGACTCGAACGTGCTT 596
   ::::: ::::: ::::: ::::: ::::: :::::
Db 170 GlySerTrpTrpGlnHisValGlnGluTrpTrpGluLeuSerHisThrHisProValLeu 189
   ::::: ::::: ::::: ::::: ::::: :::::
QY 597 TTCTCAAGTATGAACATGATCGGACCTCGTGCAGCATGGTGGAGCAGCTGGCCACA 656
   ::::: ::::: ::::: ::::: ::::: :::::
Db 190 TyrLeuPheTyrrGluAspLysGluAspProLysArgGluLeuGlnLysleLeuGlu 209
   ::::: ::::: ::::: ::::: ::::: :::::
QY 657 TTCTCGGGGTGCTGTGACAGGCCAGCTCGGAGCCCTG-----ACGGAGCAC 707
   ::::: ::::: ::::: ::::: ::::: :::::
Db 210 PheileGlyArgSerLeuProGluGluThrValAspHisleValGlnArgThrSerPhe 229
   ::::: ::::: ::::: ::::: ::::: :::::
QY 708 TGGCACCAAGTGTGGACAGTCTGCAACGCTGAGCGCTCGCCGTGGGC----- 758
   ::::: ::::: ::::: ::::: ::::: :::::
Db 230 LysGluMetLysLysAsnProMetThrAsnTyrrSerThrleProThrAlaValMetAsp 249
   ::::: ::::: ::::: ::::: ::::: :::::
QY 759 -----CGGGGAAGAGTTGGGTGTGGAGGAGCATCTTCACC 794
   ::::: ::::: ::::: ::::: ::::: :::::
Db 250 HisSerIleSerAlaPheMetArgLysGlyleThrGlyAspTrpLysSerThrPheThr 269
   ::::: ::::: ::::: ::::: ::::: :::::
QY 795 GTCTCCATGAATGAGAAGTTTGATTTGGTGTATATAACAGAAGATG 839
```

Db 270 ValAlaGlnAsnGluLeuPheGluAlaHisTyrrAlaLysLysMet 284  
RESULT 27  
A44011

adrenocortical estrogen sulfotransferase - guinea pig  
C:Species: Cavia porcellus (guinea pig)  
C:Date: 27-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A44011  
R:Oeda, T.; Lee, Y.C.; Driscoll, W.J.; Chen, H.C.; Strott, C.A.  
Mol. Endocrinol. 6, 1216-1226, 1992  
A:Title: Molecular cloning and expression of a full-length complementary DNA encoding th  
A:Reference number: A44011; MUID:93024479; PMID:1406700  
A:Accession: A44011

A>Status: preliminary  
A:Molecule type: mRNA; protein  
A:Residues: 1-296 <OED>  
A:CROSS-references: UNIPROT:P49887; UNIPARC:UPI0000136224; GB:U09552; GB:S45979; NID:949  
A>Note: sequence extracted from NCBI backbone (NCBIN:115159, NCBIP:115160)  
C:Superfamily: alcohol sulfotransferase

Alignment Scores:  
Pred. No.: 7.23e-23 Length: 296  
Score: 408.50 Matches: 94  
Percent Similarity: 52.4% Conservative: 50  
Best Local Similarity: 34.2% Mismatches: 112  
Query Match: 9.3% Indels: 19  
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x A44011 (1-296)

```
QY 72 AGTACTTTC--GAGTTCATGGCTGGCTGCCCGCTTCTGCCGCGGGAAGATGGAG 128
   ::::: ::::: ::::: ::::: ::::: :::::
Db 11 GlufyrrPheAspGluPheArgGlylleLeuLeuTyrrLysGlnPheIleLysTyrrTrpAsp 30
   ::::: ::::: ::::: ::::: ::::: :::::
QY 129 GAGATGCCCAACTTCCCGTGGCGGCCAGCAGCTGTGGATCGTCACTTACCCCAAGTCC 188
   ::::: ::::: ::::: ::::: ::::: :::::
Db 31 AsnValGluAlaPheGlnAlaArgProAspAspLeuValleAlaAlaTyrrProLysSer 50
   ::::: ::::: ::::: ::::: ::::: :::::
QY 189 GGCACCAAGTCTGCTGAGGAGTGTCTACTTGTGGAGCCGCGCTGACCCCGATGAG 248
   ::::: ::::: ::::: ::::: ::::: :::::
Db 51 GlyThrThrTrpIleSerGluValValCysMetIleTyrrAlaGluGlyAspValLysLys 70
   ::::: ::::: ::::: ::::: ::::: :::::
QY 249 ATCGGTTGATGAACATCAGCAGCAGCTCCCGCTCGGACTCCACAGCCG----- 302
   ::::: ::::: ::::: ::::: ::::: :::::
Db 71 CysArgGlnAspAlaIlePheAsnArgValProPheLeuGluCysArgAsnAspLysMet 90
   ::::: ::::: ::::: ::::: ::::: :::::
QY 303 -----GGCTGGACATCATCAAGAACTGACTCTCCCGCTCATCAAGACCACTG 356
   ::::: ::::: ::::: ::::: ::::: :::::
Db 91 MetAsnGlyValLysGlnLeuGluGluMetAsnSerProArgIlelleLysThrHisLeu 110
   ::::: ::::: ::::: ::::: ::::: :::::
QY 357 CCTACCGCTTCTGCGCTCTGACCTCCCAATGGAGACTCCAAAGTCTATATATGGCT 416
   ::::: ::::: ::::: ::::: ::::: :::::
Db 111 ProProArgLeuLeuProAlaSerPheTrpGluLysArgCysLysMetIleCysleCys 130
   ::::: ::::: ::::: ::::: ::::: :::::
QY 417 CGCAACCCCAAGATCTGTGTGTCTTATATCATGTTCCACGCTCTCTCGGACCATG 476
   ::::: ::::: ::::: ::::: ::::: :::::
Db 131 ArgAsnAlaLysAspValAlaValSerTyrrTyrrPhePheLeuMetValAlaAsnHis 150
   ::::: ::::: ::::: ::::: ::::: :::::
QY 477 AGCTACCGAGGACCTTTCAGAAATTTGCCCGAGGTTTATGAATGATGAGCTGGGCTAC 536
   ::::: ::::: ::::: ::::: ::::: :::::
Db 151 ProAspProGlySerPheProGluPheValGluLysPheMetGlnGlyGlnValProTyrr 170
   ::::: ::::: ::::: ::::: ::::: :::::
QY 537 GGCTCTGTTTGGACACGTCGAGGAGTTCGGGACCGCGCATGGACTCGAACGTGCTT 596
   ::::: ::::: ::::: ::::: ::::: :::::
Db 171 GlySerTrpTrpAspHisVallySerTrpTrpGluLysSerThrAspProArgIleLeu 190
   ::::: ::::: ::::: ::::: ::::: :::::
QY 597 TTCTCAAGTATGAACATGATCGGACCTCGGAGCTGGTGCAGCATGGTGGCCACA 656
   ::::: ::::: ::::: ::::: ::::: :::::
Db 191 PheilePheTyrrGluAspMetLysGluAspIleArgLysGluValLeuLysleulleHis 210
   ::::: ::::: ::::: ::::: ::::: :::::
QY 657 TTCTCGGGGTGCTGTGTGACAGGCCAGCTCGGAGCCCTGAGGAGCAGTCCAC--- 713
   ::::: ::::: ::::: ::::: ::::: :::::
```

Db 211 PheLeuGlyArgLysProSerGluGluLeuValAspLysIleIleLysHisThrSerPhe 230  
 QY 714 -----CAGCTGGTGACCAAGTGTGCAAC 737  
 Db 231 GlnGluMetLysAsnAsnProSerThrAsnTyrThrMetLeuProGluGluIleMetAsn 250  
 QY 738 GCTGAGGCCCTGCC--GTGGCCGGGAAGAGTTGGGCTGTGGGAAGACATCTTCACC 794  
 Db 251 GlnLysValSerProPheMetArgLysGlyIleSerGlyAspTrpLysAsnHisPheThr 270  
 QY 795 GTCTCCATGAATGAGAAGTTGACTTGGTGTGTTATAACAGAGAATG 839  
 Db 271 ValAlaLeuAsnGlnSerPheAspLysHisTyrGlnGlnMet 285  
 RESULT 28  
 JW0078  
 N:amino sulfo transferase (EC 2.8.2.3) RB1 - rabbit  
 N:Alternate names: AST-RB1  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C>Date: 17-Jun-1998 #sequence\_revision 10-Jul-1998 #text\_change 09-Jul-2004  
 C:Accession: JW0078  
 R:Yoshinari, K.; Nagata, K.; Ogino, M.; Fujita, K.; Shiraga, T.; Iwasaki, K.; Hata, T.; J. Biochem. 123, 479-486, 1998  
 A>Title: Molecular cloning and expression of an amine sulfo transferase cDNA: a new gene  
 A:Reference number: JW0078; MUID:98207028; PMID:9538231  
 A:Accession: JW0078  
 A:Molecule type: mRNA  
 A:Residues: 1-301 <YOS>  
 A:Cross-references: UNIPROT:O46640; UNIPARC:UPI0000086C35; DDBJ:D86219; NID:g2916982; PI  
 A:Experimental source: liver  
 C:Comment: This protein catalyzes 4-phenyl-1,2,3,6-tetrahydropyridine sulfation.  
 C:Superfamily: alcohol sulfo transferase  
 C:Keywords: sulfo transferase

Alignment Scores:  
 Pred. No.: 4.97e-22 Length: 301  
 Score: 397.50 Matches: 85  
 Percent Similarity: 55.4% Conservative: 68  
 Best Local Similarity: 30.8% Mismatches: 105  
 Query Match: 9.0% Indels: 18  
 DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x JW0078 (1-301)

QY 123 ATGGAGGAGATGCCAATCTCCGGTGGCCGACGAGCGTGTGGATCTCACTACCC 182  
 Db 26 LeuGluLysLeuAspAspPheGluIleAArgAspAspValPheValIleThrTyrPro 45  
 QY 183 AGTCCGGCACCAGCTTGCTGCGAGGAGTGTCTACTTGGTGACCGCCGCTGACCCC 242  
 Db 46 LysSerGlyThrValTrpThrGlnGlnIleLeuSerLeuIleTyrPheGluGlyHisArg 65  
 QY 243 GATGAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGTCTCTGGAGTACCCACAGCG 302  
 Db 66 AsnArgThrGluLysTrpAspThrLeuAspArgValProPheLeuGluTyrAsnIleArg 85  
 QY 303 GGCCTGGACATCATCAAGAACTACCTCTCCCGCCTCATCAAGAGCCACCTGCCCTAC 362  
 Db 86 LysValAsp--IleGluAsnArgProSerProArgLeuPheAlaSerHisLeuProTyr 104  
 QY 363 CGCTTCTGCTCTGACCTCCACATGGAGACTCCAGGCTCATCTATATGCTGCCAAC 422  
 Db 105 TyrLeuAlaProLysSerLeuLysAsnAsnLysAlaLysIleIleTyrValTyrArgAsn 124  
 QY 423 CCCAAGGATCTGGTGTGCTTATTATATGATTTCCACCGCTCTCTGGGACCATGAGCTAC 482  
 Db 125 ProLysAspValLeuIleSerPhePheHisPheSerAsnMetValValLysLeuGluAla 144  
 QY 483 CGAGGACCTTTCAAGAATTCTGCCGAGGTTTATGAATGATAAGTGGGCTACGGCTCC 542  
 Db 145 SerAsnThrLeuGluAsnPheMetGluLysPheLeuAspGlyLysValValGlySerIle 164  
 QY 543 TGGTTTGACACGTGACGAGATTCTGGGAGCACCCGATGGATCGAACGCTGCTTTTCTC 602

Db 165 TrpPheAspHisIleArgGlyTrpTyrGluHisLysAsnAspPheAsnIleLeuPheMet 184  
 QY 603 AAGTATGAAGACATCATCGGACCTGTGTGACGATGGTGGAGCAGCTGCCAGATTCCCTG 662  
 Db 185 MetTyrGluAspMetLysLysAspLeuArgSerSerIleLeuLysIleSerPheLeu 204  
 QY 663 GGGGTGTCTGTGACAAAGCCCGAGCTGGAAGCCCTGACGGACACTGC----- 710  
 Db 205 GluLysAspLeuSerGluGluValAspAlaIleValArgGlnAlaThrPheGluAsn 224  
 QY 711 CACCAGCTGGTGGACAGTGC-----TGCACGCTGAGGCCCTGCGCTGGGCCGG--- 761  
 Db 225 MetLysPheIleProGlnAlaAsnTyrAsnAsnIleLeuSerAsnGluIleGlyArg 244  
 QY 762 -----GGAAGAGTTGGCTGTGGAAGGACACTTTCACC 794  
 Db 245 HisAsnGluGlyAlaPheLeuArgLysGlyAlaValGlyAspTrpLysHisMetThr 264  
 QY 795 GTCTCCATGAATGAGAAGTTGACTTGGTGTATATAACAGAGATGGGAAGTGTGACCTC 854  
 Db 265 ValGluGlnSerGluArgPheAspArgIlePheGlnGluGluMetLysAspPheProLeu 284  
 QY 855 ACCTTTGACTTTTATTATATAACAGAAACAAACCACTGCATGCT 900  
 Db 285 LysPheIleTrpAspLeu-AsnAspGluAlaAsnSerAsnHisSer 299  
 RESULT 29  
 JC4531  
 N:alcohol sulfo transferase (EC 2.8.2.2) 2 - guinea pig  
 N:Alternate names: adrenal hydroxysteroid sulfo transferase 2  
 C:Species: Cavia porcellus (guinea pig)  
 C>Date: 15-Feb-1996 #sequence\_revision 19-Apr-1996 #text\_change 09-Jul-2004  
 C:Accession: JC4531; PC4120  
 R:Luu, N.X.; Driscoll, W.J.; Martin, B.M.; Strott, C.A. Biochem. Biophys. Res. Commun. 217, 1078-1086, 1995  
 A>Title: Molecular cloning and expression of a guinea pig 3-hydroxysteroid sulfo transferase  
 A:Reference number: JC4531; MUID:96125350; PMID:8554560  
 A:Accession: JC4531  
 A:Molecule type: mRNA  
 A:Residues: 1-287 <LUU>  
 A:Cross-references: UNIPROT:P52841; UNIPARC:UPI0000170746; GB:U35115; NID:g1151080; PID:N  
 A:Experimental source: adrenal gland  
 A:Accession: PC4120  
 A:Molecule type: protein  
 A:Residues: 65-73;75-83;95-103;206-215;274-287 <LU2>  
 A:Cross-references: UNIPARC:UPI000017581C; UNIPARC:UPI000017581D; UNIPARC:UPI000017581E,  
 C:Comment: This enzyme sulfonates both 3 alpha- and 3 beta-hydroxylated neutral steroids  
 C:Genetics:  
 A:Gene: gpHST2  
 A:Superfamily: alcohol sulfo transferase  
 C:Keywords: adrenal gland; sulfo transferase  
 F:249-255/Region: DNA binding #status predicted

Alignment Scores:  
 Pred. No.: 4.01e-21 Length: 287  
 Score: 385.50 Matches: 88  
 Percent Similarity: 53.4% Conservative: 55  
 Best Local Similarity: 32.8% Mismatches: 108  
 Query Match: 8.8% Indels: 17  
 DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x JC4531 (1-287)

QY 84 TTCCATGCGTGGCTGCCG-----CCCTTCTGCCGGGAAGATGGAGAGATCGCC 137  
 Db 8 PheGluGlyIleArgPheProMetValGlyPheSerProGluLeuLeuArgGluValArg 27  
 QY 138 AAC---TTCCCGGTGCGGCCCGACGCGATGTGGATCGTCACCTACCCCAAGTCGGCACC 194  
 Db 28 AspLysPheLeuValLysAspGluAspThrIleThrValThrTyrProLysSerGlyThr 47  
 QY 195 ACCTTGTCGAGGAGTGTCTACTTGTGTAGCCAGCGCGCTGACCCCGATGAGATCGGC 254



A:Accession: A54026  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-287 <LEE>  
A:Cross-references: UNIPROT:P50234; UNIPARC:UPI0000170743; GB:U066871; NID:g508212; PIDN:  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: sulfotransferase

Alignment Scores:  
Pred. No.: 1,93e-20 Length: 287  
Score: 376.50 Matches: 88  
Percent Similarity: 51.1% Conservative: 49  
Best Local Similarity: 32.8% Mismatches: 114  
Query Match: 17 Indels: 17  
DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x A54026 (1-287)

```
QY 84 TTCATGGCGTGGCGTGGCGCCCTTCGCCGGGGAAG-----ATGAGGAGATCGCC 137
Db 8 PheGluGlyIleArgPheProMetValCysPheSerProGluIleLeuArgGluValArg 27

QY 138 AAC---TTCGCCGTGGCGCCAGCAGCGTGTGGATCGTCACCTACCCCAAGTCCGCGACC 194
Db 28 AspLysPheLeuValIleAspGluAspThrIleThrValThrTyProLysSerGlyThr 47

QY 195 AGCTTGCTGACGAGGTGTCTACTTGGTGAGCCAGCGCGCTGACCCCGATGAGATCGCG 254
Db 48 AsnTrpLeuAsnGluIleValCysLeuIleLeuSerLysGlyAspProLysLeuValGln 67

QY 255 TTGATGAACATCGACGAGCGTCCGGTCTGGAGTACCCACACGCCGGCGCTGGACATC 314
Db 68 SerValProAsnTrpAspArgSerProTrpIleGluPhe---ThrGlyGlyTyrgluLeu 86

QY 315 ATCAAGGAATCGACCTCTCCCGCTCATCAAGAGCCAGCTCCCTACCGCTTCTGCGCC 374
Db 87 ValLysGlyGlnLysAspProArgValTyThrSerHisLeuProLeuHisLeuPhePro 106

QY 375 TCTGACCTCCCAATGGAGACTCCAAAGTCTATATATGGTTCGCAACCCCAAGGATCTG 434
Db 107 LysSerPhePheSerSerLysAlaLysValIleTyrcysIleArgAsnProArgAspAla 126

QY 435 GTGGTGCTTATTATCAGTTCCACCGCTCTCTGCGGACCATGAGTACCGAGGACACCTTT 494
Db 127 LeuValSerGlyTyPhePheLeuSerLysMetAsnValThrGluLysProGluThrLeu 146

QY 495 CAAGAATTCGCGGAGGTTTATGAATGATAAGCTGGGTACGGCTCCTGTTTGGAGCAC 554
Db 147 GlnGlnTyMetGluTrpPheLeuGlnGlyAsnValIleTyrglyserIrpPheGluHis 166

QY 555 GTGACGAGTTCCTGGGAGCACCGCATGGACTCGAACGTGCTTTTCTCAAGTATGAAGAC 614
Db 167 ValArgGlyTyTrpLeuSerMetArgGluMetGluAsnValLeuValSerTyrgluAsp 186

QY 615 ATGCATCGGACCTGGTACGATGTGGAGCAGCTGGCCAGATTCTCGGGGTGTCCTGT 674
Db 187 LeuIleLysAspThrArgSerThrValGluLysIleCysGlnPheLeuGlyLysLysLeu 206

QY 675 GACAGGCCAGCTGGAAGCCCTGACGCGGACACTGC-----CACCAGCTGTGGAC 725
Db 207 LysProGluGluThrAspLeuValLeuLysTySerSerPheGlnPheMetLysGluAsn 226

QY 726 CAGTCTGCAACGCTGAGCGCCCTGCCC-----GTG 755
Db 227 GluMetSerAsnPheThrLeuLeuProHisAlaTyThrThrGluGlyPheThrLeuLeu 246

QY 756 GCCCGGGAAGATTGGGTGTGGAGGACATCTTTCACCGCTCTCCATGAATGAGAAGTTT 815
Db 247 ArgLysGlyThrValGlyAspTrpLysAsnHisPheThrValAlaGlnAlaGluAlaPhe 266

QY 816 GACTTGGTGATATACAGAGATG 839
Db 267 AspLysIleTyrglnGluLysMet 274
```

RESULT 32  
JE0152

alcohol sulfotransferase (EC 2.8.2.2) - rabbit  
N/Alternate names: AS1-RB2; hydroxysteroid sulfotransferase  
C/Species: Oryctolagus cuniculus (domestic rabbit)  
C/Date: 13-Jun-1998 #sequence\_revision 10-Jul-1998 #text\_change 09-Jul-2004  
C/Accession: JE0152  
R/Yoshinari, K.; Nagata, K.; Shiraga, T.; Iwasaki, K.; Hata, T.; Ogino, M.; Ueda, R.; F.  
J. Biochem. 123, 740-746, 1998  
A/Title: Molecular cloning, expression, and enzymatic characterization of rabbit hydroxy-  
A/Reference number: JE0152; MUID:98207066; PMID:9538269  
A/Accession: JE0152  
A:Molecule type: mRNA  
A:Residues: 1-286 <YOS>  
A:Cross-references: UNIPROT:O62648; UNIPARC:UPI00000886AB; DDBJ:AB006053; NID:g3036938;  
A:Experimental source: liver  
C/Comment: This protein shows high activities to both hydroxysteroids and amines.  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: sulfotransferase

Alignment Scores:

Pred. No.: 5,06e-20 Length: 286  
Score: 371.00 Matches: 90  
Percent Similarity: 51.5% Conservative: 51  
Best Local Similarity: 32.8% Mismatches: 113  
Query Match: 8.4% Indels: 20  
DB: 2 Gaps: 6

US-10-768-158-1 (1-2419) x JE0152 (1-286)

```
QY 63 TTCGAGACCAAGTACTTCAGTTCATGGCGTGGCGCTTCGCGCGGGAAG 122
Db 8 TyrGluGlyIleAlaPheProPheValGlyPheAsn---ProGluLeuMetArgLysAla 26

QY 123 ATGAGGAGATCGCAACTTCCCGTGGCGGCCAGCAGTGTGATCGTACCTACCCC 182
Db 27 TyrGluGlu-----PheValValAsnGluAspValLeuThrValThrPhePro 43

QY 183 AAGTCGCGCACCATGCTGCTGAGGAGGTGTCTACTTGGTGAGCCAGCGCGCTGACCCC 242
Db 44 LysSerGlyThrAsnTrpLeuIleGluLeuLysLeuIleArgSerLysGlyAspAla 63

QY 243 GATGAGATCGGCTGTGATGAACATCGACGAGCAGTCCCGGTCTGTGAGTACCCACAGCGG 302
Db 64 ThrGlnIleGlnSerValProIleTrpValArgSerProTrpValGlu---ThrValSer 82

QY 303 GGCCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCGCTAC 362
Db 83 GlyTyrgluAspLeuLysThrMetGluSerProArgLeuIleSerThrHisLeuProIle 102

QY 363 CGCTTTCTGCGCTCTGACCTCCACAATGGAGACTCCAAAGTCCATCTATATGCTCGCAAC 422
Db 103 HisIlePheProLysSerIleHisThrThrLysAlaLysValIleTyLeuMetArgAsn 122

QY 423 CCCAAGATCTGTTGGTGTCTTATTATCAGTTCCACCGCTCTCTGCGGACCATGAGTAC 482
Db 123 ProArgAspValLeuValSerGlyTyTrpThrPheTrpAsnTyValLysPheValGluAsn 142

QY 483 CGAGGCACCTTTCAAGAAATTCGCGGAGGTTTATGAATGATAGCTGGGTACCGGTCC 542
Db 143 ThrLysSerLeuGlnGluTyPheGluLeuPheLeuAspGlyAsnValPheGlySer 162

QY 543 TGGTTTGAGCAGCTGACGAGGTCTCGGAGCACCGCATGGACTCGAAGCTCTTTTCTC 602
Db 163 TrpPheAspHisValHisGlyTrpLeuSerLeuArgAspGlnLysAsnPheLeuLeu 182

QY 603 AAGTATGAAGACATGCATCGGACCTGGTGCAGTGTGGAGCAGCTGGCCAGATTCTCTG 662
Db 183 SerTyrgluLeuLysGlnAsnThrArgSerThrIleGluLysIleCysHisPheLeu 202

QY 663 GGGGTGTCTGTGACAAAGGCCACAGCTG-----GAAGCC 695
Db 663 GGGGTGTCTGTGACAAAGGCCACAGCTG-----GAAGCC 695
```











Db	28	AspGluPheValIleArgAspGluAspValIleLeuThrTyrProLysSerGlyThr	47
QY	195	AGCTTGTCTCAGAGAGTGGTCTACTTTGGTGAGCCAGGGCGCTCACCCTCCGATGAGATCGGC	254
Db	48	AsnTrpLeuAlaGluIleLeuCysLeuMetHisSerLysGlyAspAlaLysTrpIleGln	67
QY	255	TTGATGAACATCAGCAGCAGCTCCGGTCTGGAGTACCCACAGCCGGCGCTGGACATC	314
Db	68	SerValProIleTrpGluArgSerProTrpValGlu--SerGluIleGlyTyrThrAla	86
QY	315	ATCAAGGAAGTCACTCTCCCGCTCATCAAGAGCCACTGCCTTACCGCTTCTTGCC	374
Db	87	LeuSerGluThrGluSerProArgLeuPheSerSerHisLeuProIleGlnLeuPhePro	106
QY	375	TCTGACTCCACAATGAGACTCCAAAGTCACTATATATGGCTCGCAACCCCAAGGATCG	434
Db	107	LysSerPhePheSerLysAlaLysValIleTyrLeuMetArgAsnProArgAspVal	126
QY	435	GTGGTGTCTTATATACAGTTCCACCGCTCTCTGGCGGACCATGACTACCGAGCACTTT	494
Db	127	LeuValSerGlyTyrPhePheTrpLysAsnMetLysPheIleLysLysProLysSerTrp	146
QY	495	CAAGAATCTCGCGGAGTTTATGAATGATAAGCTGGGCTACGGCTCCTGGTTTGAGCAC	554
Db	147	GluGluTyrPheGluTrpPheCysGlnGlyThrValLeuTyrGlySerTrpPheAspHis	166
QY	555	GTGAGGAGTTCGGAGCACCAGCATGGACTCGAAGTGTCTTTCTCAAGTATGAAGAC	614
Db	167	IleHisGlyTrpMetProMetArgGluGluLysAsnPheLeuLeuLeuSerTyrGluGlu	186
QY	615	ATGCATCGGACCTGGTACGATGGTGAGCAGCTGGCCAGATTTCTGGGGGTGTCCTGT	674
Db	187	LeuLysGlnAspThrGlyArgThrIleGluLysIleCysGlnPheLeuGlyLysThrLeu	206
QY	675	GACAAGGCCACGTGGAAGCCCTGACGGAGCACTGC-----CACCAGCTGTGGAC	725
Db	207	GluProGluGluLeuAsnLeuIleLysAsnSerSerPheGlnSerMetLysGluAsn	226
QY	726	CAGTGCTGCAACGCTGAGGCCCTGCCGTGGGC-----	758
Db	227	LysMetSerAsnTyrSerLeuLeuSerValAspTyrValValAspLysAlaGlnLeuLeu	246
QY	759	---CGGGAAGATTGGGCTGTGGAGGACATCTTCACCGTCTCCATGATGAGAGTTT	815
Db	247	ArgLysGlyValSerGlyAspTrpLysAsnHisPheThrValAlaGlnAlaGluAspPhe	266
QY	816	GACTTGGTGATAACAGAGATG	839
Db	267	AspLysLeuPheGlnGluLysMet	274
RESULT 37			
152849			
alcohol sulfotransferase (EC 2.8.2.2) - rat			
N:Alternate names: hydroxysteroid sulfotransferase			
C:Species: Rattus norvegicus (Norway rat)			
C>Date: 12-Aug-1996 #sequence revision 12-Aug-1996 #text_change 09-Jul-2004			
C:Accession: I52849; I65759			
R:Watabe, T.; Ogura, K.; Satsukawa, M.; Okuda, H.; Hiratsuka, A.			
Chem. Biol. Interact. 92, 87-105, 1994			
A:Title: Molecular cloning and functions of rat liver hydroxysteroid sulfotransferases c			
A:Reference number: I52849; MUID:94306585; PMID:8033273			
A:Accession: I52849			
A:Status: preliminary; translated from GB/EMBL/DBJ			
A:Molecule type: mRNA			
A:Residues: 1-284 <RES>			
A:Cross-references: UNIPROT:P15709; UNIPARC:UPI0000170BD9; GB:D14987; NID:9303797; PIDN:			
C:Superfamily: alcohol sulfotransferase			
C:Keywords: sulfotransferase			
Alignment Scores:			
Pred. No.:	1,08e-18	Length:	284
Score:	353.50	Matches:	89
Percent Similarity:	52.2%	Conservative:	54

Best Local Similarity: 32.5%			
Query Match: 8.0%			
DB:	2	Mismatches:	108
		Indels:	23
		Gaps:	7
US-10-768-158-1 (1-2419) x I52849 (1-284)			
QY	75	TACTTCGAGTTCATATGGCTGCGGCTGCGCCCTTCTGCGCGGGAAG-----ATGGAG	128
Db	4	TyrThrTrpPheGluGlyIleProPheHisAlaPheGlyIleSerLysGluThrLeuGln	23
QY	129	GAGATCGCCAAC---TTCCTGGTGGCGCCAGCAGCTGTGGATCGTCACTACCCCAAG	185
Db	24	AsnValCysAsnLysPheValValLysGluGluAspLeuLeuAlaTyrProLys	43
QY	186	TCCGCCACCAAGCTTCTGCGAGGAGTGGTCTACTTTGGTGAGCCAGGGCGCTCACCCCGAT	245
Db	44	SerGlyThrAsnTrpLeuIleGluIleValCysLeuIleGlnThrLysGlyAspProLys	63
QY	246	GAGATCGGCTTGATGAACATGACGAGCAGCTCCGGTCTCTGGAGTACCCACAGCGGGC	305
Db	64	TrpIleGlnSerValThrIleTrpAspArgSerProTrpIleGlu---ThrAspValGly	82
QY	306	CTGGACATCATCAAGGAAGTCACTCTCCCGCTCATCAAGAGCCACTGCGCTACCCG	365
Db	83	TyrAspIleLeuIleLysLysGlyProArgLeuMetThrSerHisLeuProMetHis	102
QY	366	TTTCTGCCCTCTGACCTCCACAATGGAGACTCCAAGTGCATATATGGCTCGCAACCCC	425
Db	103	LeuPheSerLysSerLeuPheSerSerLysAlaLysValIleTyrLeuIleArgAsnPro	122
QY	426	AAGATCTGGTGGTCTTATATCAGTTTCCACCCGCTCT-----CTGCGGACCATG	476
Db	123	ArgAspValLeuValSerGlyTyrTyrPheTrpGlyAsnSerThrLeuValLysPro	142
QY	477	AGCTACCGAGCAGCCTTTCAGAAATTCGCGGAGGTTTATCAATGATAAGTGGGCTAC	536
Db	143	AspSerLeuGlyThrTyrValGluTrp-----PheLeuLysGlyAsnValLeuTyr	159
QY	537	GGCTCTCTGTTGACGAGTTCGAGGAGTTCGGGAGCACCGCATGGACTCGAACGTGCTT	596
Db	160	GlySerTrpPheGluHisIleArgAlaTrpLeuSerMetArgGluTrpAspAsnPheLeu	179
QY	597	TTTCTCAAGTATGAAGATGATGATCGGAGCCTGTGTGACGATGGTGGAGCAGCTGGCCAGA	656
Db	180	LeuLeuTyrTyrGluAspMetLysLysAspThrMetGlyThrIleLysLysIleCysAsp	199
QY	657	TTCTGGGGGTGCTCTGTGACAAAGCCAGCTGGAAGCCCTGACGGAGCATGCGCACAG	716
Db	200	PheLeuGlyLysLysLeuGluProAspGluLeuAspLeuValLeuLysTyrSerSerPhe	219
QY	717	CTGGTG-----GACCAGTGTGCAACCTGAGGCCCTGGCC-----	752
Db	220	GlnValMetLysGluAsnAspMetSerAsnTyrSerLeuLeuMetLysLysSerIlePhe	239
QY	753	-----GTGGCGCGGGAAGAGTTGGGCTGTGGGAAGACATCTTCCACGTC	797
Db	240	ThrGlyIleGlyLeuMetArgLysGlyThrIleGlyAspTrpLysAsnHisPheThrVal	259
QY	798	TCCATGATGAGAAGTTGACTTGGTGTATATAACAGAGATG	839
Db	260	SerGlnAlaGluAlaPheAspLysValPheGlnGluLysMet	273
RESULT 38			
A33569			
alcohol sulfotransferase (EC 2.8.2.2) - rat			
C:Species: Rattus norvegicus (Norway rat)			
C>Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 09-Jul-2004			
C:Accession: A33569			
R:Ogura, K.; Kajita, J.; Narihata, H.; Watabe, T.; Ozawa, S.; Nagata, K.; Yamazoe, Y.; K			
Biochem. Biophys. Res. Commun. 165, 168-174, 1989			
A:Title: Cloning and sequence analysis of a rat liver cDNA encoding hydroxysteroid sulfo			
A:Reference number: A33569; MUID:90073705; PMID:2590219			
A:Accession: A33569			

A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-284 <OG>  
A;Cross-references: UN  
C;Superfamily: alcohol  
C;Keywords: sulfotrans

**Alignment Scores:**

Alignment scores:	
Pred. No.:	1,53e-18
Score:	351.50
Percent Similarity:	52.7%
Best Local Similarity:	32.7%
Query Match:	8.0%
DB:	2
Gaps:	25
Indels:	8
Mismatches:	105
Conservative:	55
Matches:	90
Length:	284

US-10-76A-158-1 (1-2419) x A33569 (1-284)

QY	75	TACTTCGAGTTCCATGCGGTGCGGTGCGCGCCCTTCTGCGCGGGAAG-----ATGAG	128
Db	4	TyrThrTrpPheGluGlyIleProPheHisAlaPheGlyIleSerLysGluThrLeuGln	23
QY	129	GAGATCGCCAAAC---TTCCCGGTGCGGCCCGACGCGTGTGATCGTCACTACCCCAAG	185
Db	24	AsnValCysAsnLysPheValValLysAspGluAspLeuLeuLeuAlaTyrProLys	43
QY	186	TCCGGCACCGACTGCTGCGAGGAGTGGTCTACTTTGGTGAGCCAGCGGCGCTGACCCCGAT	245
Db	44	SerGlyThrAsnTrpLeuLeuLeuValCysLeuLeuGlnThrLysGlyAspProLys	63
QY	246	GAGATCGGCTTGATGAACATCAGCAGCAGCTCCCGGTCTCTGGATGCCACAGCGGGC	305
Db	64	TrpIleGlnSerValThrIleTrpAspArgSerProTrpIleGlu---ThrAspValGly	82
QY	306	CTGGACATCATCAAGGAACGTACCTCTCCCGCCCTCATCAAGAGCCACCTGCGCTACCGC	365
Db	83	TyrAspIleLeuIleLysLysGlyProArgLeuMetThrSerHisLeuProMetHis	102
QY	366	TTTCTGCGCTCTGACCTCCACATCGAGACTCCAAGGTCACTCTATGCTCGCAACCCC	425
Db	103	LeuPheSerLysSerLeuPheSerLysAlaLysValIleTyrLeuValArgAsnPro	122
QY	426	AAGGATCTGGTGGTCTTATTATCATGTTCCACCGCTCTCTCGGACCATGAGCTACCGA	485
Db	123	ArgAspValLeuValSerGlyTyrTyrPheTrpGlyAsn---SerThrLeuAlaLysLys	141
QY	486	-----GGCACCTTCAAGAATCTCTCGCGAGGCTTATGAATGATGAAGCTGGGC	533
Db	142	ProAspSerLeuGlyThrTyrValGluTrp-----PheLeuLysGlyAsnValLeu	158
QY	534	TACGGCTCTGTTGAGACGCTGTCAGGAGTCTTGGGACACCGCATGGAAGCTGCAACGTTG	593
Db	159	TyrGlySerTrpPheGluHisIleArgAlaTrpLeuSerMetGlnGluTrpAspAsnPhe	178
QY	594	CTTTTCTCAAGTATGAACACATGCATCGGACCTGGTGACGATGCTGGAGCAGCTGGCC	653
Db	179	LeuLeuLeuTyrTyrGluAspMetLysLysAspThrMetGlyThrIleLysLysIleCys	198
QY	654	AGATTCTCGGGGTGTCCTGTGCACAGGCCACGCTGGAAGCCCTGACGGAGCACTCCAC	713
Db	199	AspPheLeuGlyLysLysLeuGluProAspGluLeuAspLeuValLeuLysTyrSerSer	218
QY	714	CAGCTGGTG-----GACCATGCTGCAACGCTGAGCCCTGAGCCCTGCGCC	752
Db	219	PheGlnValMetLysGluAsnAspMetSerAsnTyrSerLeuLeuMetLysLysSerIle	238
QY	753	-----GTGGCCCGGGGAAGTTGGGTGGGAGGACATCTTCACC	794
Db	239	PheThrGlyThrClyLeuMetArgLysGlyThrValGlyAspTrpLysAsnHisPheThr	258
QY	795	GTCTCCATGAATGAGAGGTTTGACTTGGTGTATATAACAGAGATG	839
Db	259	ValSerGlnAlaGluAlaPheAspLysValPheGlnGluLysMet	273

RESULT 39

A26136  
sensequence marker protein 2, hepatic - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 05-Oct-1988 #sequence\_revision 05-Oct-1988 #text\_change 09-Jul-2004  
C:Accession: A26136  
E:Chatterjee, B.; Majumdar, D.; Ozbilen, O.; Murty, C.V.R.; Roy, A.K.  
J. Biol. Chem. 262, 822-825, 1987  
A:Title: Molecular cloning and characterization of cDNA for androgen-repressible rat liv  
A:Reference number: A26136; MUID:87109186; PMID:3805009  
A:Accession: A26136  
A:Molecule type: mRNA  
A:Residues: 1-282 <CHA>  
A:Cross-references: UNIPROT:P07631; UNIPARC:UPI000017581B  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: liver

Alignment Scores:

Alignment Scores:		
Pred. NO.:	1.62e-17	282
Score:	338.00	89
Percent Similarity:	48.9%	47
Best Local Similarity:	32.0%	112
Query Match:	7.7%	30
DR:	2	6
	Length:	
	Matches:	
	Conservative:	
	Mismatches:	
	Indels:	
	Gaps:	

IIS-10-768-158-1 (1-2419) x A26136 (1-282)

QY	69	AGCAAGTACTTCGAGTTCCATGCGGTGGCGTGCCG-----CCTTTCTGCCGGGGGAAG	122
Db	3	SerAspTyrAsnTrpPheGluGlyIleProPheProAlaIleSerTyrGlnArgGluIle	22
QY	123	ATGAGAGGATCGCCAAC---TTCCGGTGGCGCCACGACGAGCTGTGGATCGTCAACCTAC	179
Db	23	LeuGluAspIleArgAsnLysPheValVallysGluAspLeuLeuIleLeuThrTyr	42
QY	180	CCCAAAGTCCCGGCACAGCTCTCTCGAGGAGGTGTCTACTTGGTGAGCCAGCGGCGCTCAC	239
Db	43	ProLysSerGlyThrAsnTrpLeuAsnGluIleValCysLeuIleGlnThrLysGlyAsp	62
QY	240	CCCGATGAGATC-----GGCTTGATGAATCATCGACGACGAGCTCCCCGTC	284
Db	63	ProLysTrpIleGlnSerCysProPheGlyThrValTyrProAspGluIleGluTrpIle	82
QY	285	CTGGAGTACCACACAGCCGGGCTGGACATCATCAAGGAACCTCACTCTCCCGCCTCATC	344
Db	83	PheArgAsnAsnHisGlyGly-----ProargLeuIle	93
QY	345	AAGAGCCACCTGCCCTACCGCTTCTGCGCTCTGACCTCCACAATGGAGACTCCCAAGTCT	404
Db	94	ThrSerHisLeuProIleHisLeuPheSerLysSerPhePheSerSerLysAlaLysAla	113
QY	405	ATCTATATGGTTCGCAACCCCAAGGATCTGGTGTCTATTATCAAGTTCCACCGCTCT	464
Db	114	IleTyrLeuMetArgAsnProArgAspIleLeuValSerGlyTyrPhePheTrpGlyAsn	133
QY	465	CTCGGGACCATGAGCTACCGAGGCACCTTTCNAGAATTCGCGGAGGTTTTATGAATGAT	524
Db	134	ThrAsnLeuValLysAsnProGlySerLeuGlyThrTyrPheGluTrpPheLeuGlnGly	153
QY	525	AAGCTGGGCTACGGCTCTCGGTTTCGACGACGTGCAGGAGTCTCGGAGCACCGCATGGAC	584
Db	154	AsnValLeuPheGlySerTrpPheGluHisValArgGlyTrpLeuSerMetArgGluTrp	173
QY	585	TGAAACGTGCTTTTCTCAAGTATGAAGACATGCATTCGGGACCTGGTAGCATGTGGAG	644
Db	174	AspAsnPheLeuValLeuTyrTyrgluAspMetLysLysAspThrMetGlyThrIleLys	193
QY	645	CAGCTGGCCAGATTTCTGGGGGTCTCCTGTACAGGCCACAGCTGGAAGCCCTGACGGAG	704
Db	194	LysIleCysAspPheLeuGlyLysAsnLeuGlyProAspGluLeuAspLeuLeuLys	213
QY	705	CACTGC-----CACCAAGTCTGGTGGACCAG	728

Db 214 TyrSerSerPheGlnAlaMetLysGluAenAsnMetSerAsnTyrSerLeuIleLysGlu 233

QY 729 TGCTGCAACGCTGAGGCCCTGCCCGTGGCCGG---GGAAGAGTTGGGTGTGGGAAGGAC 785

Db 234 AspProIleLeuThrGlyLeuLysLeuMetArgLysGlyThrThrGlyAspTrpLysAsn 253

QY 786 ATCTTCACCGTCTCCATGAATGAGAAGTTTGACTTGTTATATAACAGAGATG 839

Db 254 HisPheThrValAlaGlnAlaGluAlaPheAspLysValPheGlnGluLysMet 271

RESULT 40

B40216

Flavonol 3'-sulfotransferase - Flaveria chloraefolia

C:Species: Flaveria chloraefolia

C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C:Accession: B40216

R:Varin, L.; Deluca, V.; Ibrahim, R.K.; Brissson, N.

Proc. Natl. Acad. Sci. U.S.A. 89, 1286-1290, 1992

A:Title: Molecular characterization of two plant flavonol sulfotransferases.

A:Reference number: B40216; MUID:92159034; PMID:1741382

A:Accession: B40216

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-311 <VAR>

A:Cross-references: UNIPROT:P52836; UNIPARC:UPI000012A3FC; GB:M84135; NID:g168166; PID:g

A:Experimental source: terminal bud

A:Note: sequence extracted from NCBI backbone (NCBIP:82218)

C:Superfamily: alcohol sulfotransferase

Alignment Scores:

Pred. No.: 7,53e-15 Length: 311

Score: 303.00 Matches: 82

Percent Similarity: 47.9% Conservative: 57

Best Local Similarity: 28.3% Mismatches: 95

Query Match: 6.9% Indels: 56

DB: 2 Gaps: 11

US-10-768-158-1 (1-2419) x B40216 (1-311)

QY 108 TTCTGCCGCGGAAGATGAGGAGATCGCCAACTTCCCGTGGCGCCAGCGAGTGTGG 167

Db 34 PheLeuGluGlyArgIleLeuSerGluGlnLysPheLysAlaHisProAsnAspValPhe 53

QY 168 ATCGTCACCTACCCCAAGTCCGACACGAGTGTCTCGCAGGAGGTGTCTACTTGTGTGAGC 227

Db 54 LeuAlaSerTyrProLysSerGlyThrThrTrpLeuLysAlaTrpIleCysIleIleThr 73

QY 228 CAGGGCGCTGACCCCGATGAGTCGCG---TTGATGAACATC-----GACGAGCAGCTC 278

Db 74 ArgGluLysPheAspAspSerThrSerProLeuLeuThrThrMetProHisAspCysIle 93

QY 279 CCGGTCTCTGGAGTACCCACAGCGCGCGCTGGACATCATCAAGGAA----- 323

Db 94 ProLeuLeuGlu-----LysAspLeuGluLysIleGlnGluAsnGlnArgAsnSer 110

QY 324 CTGACCTCTCCCGCTCATCAAGACCCACTGCTCCCTTCTCCGCTTCTGCTGTGACCTC 383

Db 111 LeuTyrThrPro-----IleSerThrHisPheHisTyrLysSerLeuProGluSerAla 128

QY 384 CACATGGAGACTCCCAAGTCACTATATGGTTCGCAACCCAGGATCTGTGTGTCT 443

Db 129 ArgThrSerAsnCysLysIleValTyrIleTyrArgAsnMetLysAspValIleValSer 148

QY 444 TATTATCAGTTCCACCGCTCTCTCGGACCATGAGC-----TACCGAGGC 488

Db 149 TyrTyrHisPheLeuArgGlnIleValLysLeuSerValGluAlaProPheGluGlu 168

QY 489 ACCTTTCAAGAAATTCGCGGAGGTTTATCAATGATAAGCTGGGCTACGGCTCTCGGTTT 548

Db 169 AlaPheAspGluPheCysGlnGlyIleSerSer-----CysGlyProTyrTrp 184

QY 549 GAGCAGCTCAGGAGTTCTGGGAGCACCGCATGGACTCGAAC-----GTGCTTTTCTC 602

Db 185 GluHisIleLysGlyTyrTrpLysAlaSerLeuGluLysProGluIlePheLeuPheLeu 204

QY 603 AAGTATGAGACATGATCGGACCTGTGTGACGATGGTGGAGCAGCTGGCCAGATTCTCTG 662

Db 205 LysTyrGluAspMetLysLysAspProValProSerValLysLysLeuAlaAspPheIle 224

QY 663 GGG-----GTGTCC 671

Db 225 GlyHisProPheThrProLysGluGluAlaGlyValIleGluAspIleValLysLeu 244

QY 672 TGTGACAAGCCGACGTGGAAGCCCTG-----ACGGAG 704

Db 245 CysSerPheGluLysLeuSerSerLeuGluValAsnLysSerGlyMethHisArgProGlu 264

QY 705 CACTGCCACACGCTGTGGACGACGCTGTCAACGCTGAGCGCCCTGCCCCGCGCGGA 764

Db 265 GluAlaHisSerIleGluAsnArg-----LeuTyrPheArgLysGly 278

QY 765 AGAGTTGGCTGTGGAGACATCTTACCCTCTCCATCAATGAGAGTTTACTTGGTGTG 824

Db 279 LysAspGlyAspTrpLysAsnTyrPheThrAspGluMetThrGlnLysIleAspLysLeu 298

QY 825 TATAAACAGAGATGGGAAAGTGTGACCTC 854

Db 299 IleAspGluLysLeuGlyAlaThrGlyLeu 308

RESULT 41

T47448

sulfotransferase-like protein - Arabidopsis thaliana

N:Alternate names: protein T14D3.20

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004

C:Accession: T47448

R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unsel, M.; Mewes, H.W.; Lemcke, K.

submitted to the Protein Sequence Database, February 2000

A:Reference number: Z24467

A:Accession: T47448

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-329 <JUN>

A:Cross-references: UNIPROT:O9M1V1; UNIPARC:UPI000009DA7F; EMBL:AL138649

A:Experimental source: cultivar Columbia; BAC clone T14D3

C:Genetics:

A:Map position: 3

A:Note: T14D3.20

C:Superfamily: alcohol sulfotransferase

Alignment Scores:

Pred. No.: 9,08e-15 Length: 329

Score: 302.00 Matches: 89

Percent Similarity: 45.2% Conservative: 61

Best Local Similarity: 26.8% Mismatches: 108

Query Match: 6.9% Indels: 74

DB: 2 Gaps: 15

US-10-768-158-1 (1-2419) x T47448 (1-329)

QY 27 GAGAGCGAGCGCGAGACCCCGCGGAGTTCGAGAGCAAGTACTTCGAGTTC 86

Db 18 GluSerLysThrLeuIleSerSerLeuProSerAspLysAsnSerThr----- 33

QY 87 CATGCGTGGCGCTG-----CCGCGCTTCTGCGCGGG 119

Db 34 ---GlyValAsnValCysLysTyrGlnGlyCysTyrTrpTyrThrProIleLeuGlnGly 52

QY 120 AGATGGAGGAGATCGCAACTTCCGCTGGCGCCGCGAGCTGTGGTCTGCACCTAC 179

Db 53 ValLeuAsnPheGlnLysAsnPheLysProGlnAspThrAspIleValAlaSerPhe 72

QY 180 CCCAAGTCCGCGCAGCAGCTTGTCTGAGGAGGTGGTCTAC---TTGCTGAGCCAGCGCGCT 236

Db 73 ProLysCysGlyThrThrTrpLeuLysAlaLeuThrPheAlaLeuValArgArgSerLys 92

```
QY 237 GACCCCGATGATCGGCTTGATGAACATCGACGAGCAGCTCCCGTCTCTG---CAGTAC 293
Db |||
93 HisProSerHis-----AspAspHisHisProLeuLeuSerAspAsn 106
QY 294 CCA-----CAGCCGGCCTG-----GAC 311
Db |||
107 ProHisValLeuSerProSerLeuGluMetTyrLeuTyrLeuCysSerGluAsnProAsp 126
QY 312 ATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTCTG 371
Db |||
127 LeuThrLysPheSerSerSerArgLeuPheSerHisMetProSerHisThrLeu 146
QY 372 CCTCTGACCTCCAAAGGAGACTCCAAAGTCATCTATATGCTCGCAACCCCAAGGAT 431
Db |||
147 GlnGluGlyLeuLysGlySerThrCysLysIleValTyrMetSerArgAsnValLysAsp 166
QY 432 CTGGTGGTCTCTATTATCATCAGTTCCACCGCTCTCTCGGAGC-----ATG 476
Db |||
167 ThrLeuValSerTyrTrpHisPheCysLysLysGlnThrAspAsnIleIleSer 186
QY 477 AGCTACCGAGGACCTTTCAAGATTCTCCGG-----AGGTTTATGAATGATAAGCTG 530
Db |||
187 SerValGluAspThrPheGluMetPheCysArgGlyValAsnPhe-----201
QY 531 GGCTACGGTCTCTGTTTGAACAGCTGACGAGGATCTGGAGCACCGCATGGAC-----584
Db |||
202 ---PheGlyProPheTrpAspHisValLeuSerTyrTrpArgLysSerLeuGluAspPro 220
QY 585 TCGAAGCTGCTTTTCTCAAGTATGAAGACATCATCGGAGCTCTGGAGCACCGCATGGAG 644
Db |||
221 AsnHisValLeuPheMetLysPheGluGluMetLysGluGluProArgGluGlnIleLys 240
QY 645 CAGCTGGCCAGATTCTCGGGGTGCTCTGTGACAGGCCAGCCAGCTGGAAGCC---CTGACG 701
Db |||
241 ArgLeuAlaGluPheLeuGlyCysLeuPheThrLysGluGluGluSerGlyLeuVal 260
QY 702 GAGCACTGCCACGAGTGTGTGACCAAGTCTGC-----AACGCTGAGCCCTGCCGCTG 755
Db |||
261 Asp-----GluIleIleAspLeuCysSerLeuArgAsnLeuSerLeuGluIle 277
QY 756 GGC-----CGGGA 764
Db |||
278 AsnLysThrGlyLysLeuHisSerThrGlyArgGluAsnLysThrPheArgLysGly 297
QY 765 AGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCCATGAATGAGAAGTTGACTGGTG 824
Db |||
298 GluValGlyAspTrpLysAsnTyrLeuThrProGluMetGluAsnLysIleAspMetIle 317
QY 825 TATAACAGAAAGATGGGAAGTGTGACCTCACGTTT 860
Db |||
318 IleGlnGluLysLeuGlnAsnSerGlyLeuLysPhe 329

RESULT 42
B84452
probable steroid sulfotransferase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: B84452
R.; Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: B84452
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-324 <STO>
A:Cross-references: UNIPROT:Q9ZPQ5; UNIPARC:UPI000000A21B0; GB:A8002093; NID:g4406769; PI
A:Gene: At2g03770
A:Map position: 2
C:Superfamily: alcohol sulfotransferase
```

```
Alignment Scores:
Pred. No.: 1.99e-14 Length: 324
Score: 297.50 Matches: 84
Percent Similarity: 43.3% Conservative: 57
Best Local Similarity: 25.8% Mismatches: 124
Query Match: 6.8% Indels: 61
DB: 2 Gaps: 10

US-10-768-158-1 (1-2419) x B84452 (1-324)
QY 21 ATGGCGGAGAGCGAG-----GCCGAGACCCCGACGACC 53
Db |||
12 LeuValGluSerGluLeuValGlnGluCysGluGluLeuLeuSerSerLeuProArgAsp 31
QY 54 CCGGGGAGTTTCGAGAGCAAGTACTTCAGTTCCATTCGCGTCCGCGCTCCGCCCC---TTC 110
Db |||
32 ArgSerValPheAlaGluTyrLeuTyrGlnTyrPheTrpTyrProProAsnLeu 51
QY 111 TGCCCGGGAAGATGAGAGATCGCAACTTCCCGTCCGCCGAGCGAGCTGTGGATC 170
Db |||
52 LeuGluGlyValLeuTyrSerGlnLysHisPheGlnAlaArgAspSerAspIleValLeu 71
QY 171 GTCACCTACCCCAAGTCCCGCACCAGCTTGTGCGAGAGGTGCTACTTGTGTGAGCCAG 230
Db |||
72 AlaSerIleProLysSerGlyThrThrTrpLeuLysSerLeuValPhe-----87
QY 231 GCGCTGACCCCGATGAGATCGGCTTGATGAACATCGACGAG---CAGCTCCCGTCTCTG 287
Db |||
88 -----AlaLeuIleHisArgGlnGluPheGlnThrProLeuVal 100
QY 288 GAGTACCACAG-----CCGGGCGCTGGAC 311
Db |||
101 SerHisProLeuLeuAspAsnAsnProHisThrLeuValThrPheIleGluGlyPheHis 120
QY 312 ATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTTACCGCTTCTG 371
Db |||
121 LeuHisThrGlnAspThrSerProArgIlePheSerThrHisIleProValGlySerLeu 140
QY 372 CCTCTGACCTCCCAATGAGACTCCCAAGGTCTATATGCTCGCAACCCCAAGGAT 431
Db |||
141 ProGluSerValLysAspSerSerCysLysValValTyrCysArgAsnProLysAsp 160
QY 432 CTGGTGGTCTTATATATCAGTTCCACCGCTCTCTGCGGACCATGAGCTACCGAGC---488
Db |||
161 AlaPheValSerLeuThrHisPheMetLysAsnLeuIleValLysGluMetValGlyCys 180
QY 489 ACCTTTCAAGAAATTCGCGGAGGTTTATGAATGATGAAGCTGGCTACCGCTCTCTGTTT 548
Db |||
181 ThrMetGluGluMetValArgPhePheCysArgGlySerSerIleTyrGlyProPheTrp 200
QY 549 GAGCAGCTGACGAGTTCGGAGACACCGCATGGACTCG-----AACGTGCTTTTCTC 602
Db |||
201 AspHisValLeuGlnTyrTrpLysGluSerArgGluAsnProLysLysValMetPheVal 220
QY 603 AGATATGAAGACATGATCGGACCTGGTGCAGTGTGAGCAGCTGCGAGCAGCTGCGCATCTCCTG 662
Db |||
221 MetTyrGluGluMetArgGluGlnProGlnGluTrpValMetArgIleAlaGluPheLeu 240
QY 663 GGGGTGCTGTGTGACAAAGCCAGCTGGAA---GCCCTGACGAGGACACTGCCACCAAGCTG 719
Db |||
241 GlyTyrSerPheThrGluGluGluIleGluAenglyValLeuGluAspIleIleLysLeu 260
QY 720 GTGACCAAGTGTGCAACCGCTGAGGCCCTGCC---752
Db |||
261 -----CysSerLeuGluAsnLeuSerLysLeuValAsnGluLysGlyLys 276
QY 753 -----GTGGCGCGGGAAGAGTGGCTGTGGAAG 782
Db |||
277 LeuLeuAsnGlyMetGluThrLysAlaPheArgLysGlyGluIleGlyLysTrpArg 296
QY 783 GACATCTTCCCGTCTCCCATGAATGAGAAGTTTGACTTGTGTATTAACAAGAGATGGGA 842
Db |||
```

Db 297 AspThrLeuThrProLeuLeuAlaGluGluIleAspLysThrThrLysGluLysLeuIle 316  
QY 843 AAGTGTGACCTCACGTTT 860  
Db 317 GlySerAspPheArgPhe 322  
RESULT 43  
T47447  
sulfotransferase-like protein - Arabidopsis thaliana  
N:Alternate names: protein T14D3.10  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C:Accession: T47447  
R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Lemcke, K.  
submitted to the Protein Sequence Database, February 2000  
A:Reference number: Z24467  
A:Accession: T47447  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-323 <J0R>  
A:CROSS-references: UNIPROT:Q9M1V2; UNIPARC:UPI00000A990C; EMBL:AL138649  
A:Experimental source: cultivar Columbia; BAC clone T14D3  
C:Genetics:  
A:Map position: 3  
A:Note: T14D3.10  
C:Superfamily: alcohol sulfotransferase  
Alignment Scores:  
Pred. No.: 1-93e-13 Length: 323  
Score: 284.50 Matches: 74  
Percent Similarity: 45.7% Conservative: 58  
Best Local Similarity: 25.6% Mismatches: 90  
Query Match: 6.5% Indels: 67  
DB: 2 Gaps: 11  
US-10-768-158-1 (1-2419) x T47447 (1-323)  
QY 138 AACTTCCGGTGGGCCCCAGCGAGTGTGTGATCGTCACTACCCCAAGTCCGCGCACGAC 197  
Db 54 SerPheLysProGlnAspThrAspIleValAlaSerPheProLysCysGlyThrThr 73  
QY 198 TTGCTGCAGGAGTGTCTAC---TTGGTGAGCCAGGCGCTGACCCCGATGATGATCGGC 254  
Db 74 TrpLeuLysAlaLeuThrPheAlaLeuLeuHisArgSerLysGlnProSerHis----- 91  
QY 255 TTGATGAACATCGACGAGCTCCCGGTCCTG----- 287  
Db 92 -----AspAspHisProLeuLeuSerAsnAsnProHisValLeuValPro 107  
QY 288 -----GAGTACCCACAGCCGCGGCTGGACATCATC 317  
Db 108 TyrPheGluIleAspLeuThrLeuArgSerGluAsnPro-----AspLeuThr 123  
QY 318 AGGAACTGACCTCCCGCTCATCAGACGACCTGCCCTTACCGCTTTCGCCCTCT 377  
Db 124 LysPheSerSerProArgLysPheSerThrHisValProSerHisThrLysGlnGlu 143  
QY 378 GACCTCCCAATCGAGACTCCAGGTCACTATATCGCTCGCAACCCAGGATCTGGNG 437  
Db 144 GlyLeuLysGlySerThrCysLysIleValTyrLysArgAsnValLysAspThrLeu 163  
QY 438 GTGCTTATTATCAGTTCACCGCTCTCTCGCGGACC-----ATGAGCTAC 482  
Db 164 ValSerTyrTrpHisPheThrLysLysGlnThrAspGluLysIleIleSerSerPhe 183  
QY 483 CGAGGACCTTTCAAGAAATTCGCGGAGGTTTATGAATGATAGCTGGGC----- 533  
Db 184 GluAspThrPheGluMetPheCysArg-----GlyValSerIle 196  
QY 534 TAGCGCTCTGGTTTTCAGACGTCGAGGAGTTCTGGGACCGCGATGAC-----TCG 587  
Db 197 PheGlyProPheTrpAspHisValLeuSerTyrTrpArgGlySerLeuGluAspProAsn 216

QY 588 AACGTGCTTTTCTCAAGTATGAACATGTCGCGACCTCGTGACGATGCTGGAGCAG 647  
Db 217 HisValLeuPheMetLysPheGluGluMetLysAlaGluProArgAspGlnIleLysLys 236  
QY 648 CTGGCCAGATTCTCTGGGGTGTCTGTGACAAAGGCCAGCTGAAAGCCCTGACGGAGC 707  
Db 237 PheAlaGluPheLeuGly-----CysProPheThrLysGluGluGluSerGlySer 254  
QY 708 TGCCACCACTGGTGGACAGCTGCTGC-----AACGTGAGGCCCTGCGCTGGGC--- 758  
Db 255 ValAspGluIleAspLeuCysSerLeuArgAsnLeuSerSerLeuGluIleAsnLys 274  
QY 759 -----CGGGGAAGCTTGG 773  
Db 275 ThrGlyLysLeuAsnSerGlyArgGluAsnLysMetPhePheArgLysGlyGluValGly 294  
QY 774 CTGTGGAAGGACATCTTCACCGTCTCCATGAATGAGAACTTTGACTTGGTGTATAAACAG 833  
Db 295 AspTrpLysAsnTyrLeuThrProGluMetGluAsnLysIleAspMetIleleGlnGlu 314  
QY 834 AAGATGGGAAGTGTGACCTCACGTTT 860  
Db 315 LysLeuGlnAsnSerGlyLeuLysPhe 323  
RESULT 44  
T07833  
probable steroid sulfotransferase (EC 2.8.2.15) 3 - rape  
C:Species: Brassica napus (rape)  
C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 09-Jul-2004  
C:Accession: T07833  
R:Richard, M.; Nicolle, L.; Varin, L.  
submitted to the EMBL Data Library, April 1997  
A:Reference number: Z16161  
A:Accession: T07833  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-325 <R1C>  
A:CROSS-references: UNIPROT:O82410; UNIPARC:UPI00000AA137; EMBL:AF000307; NID:g3420007;  
C:Genetics:  
A:Gene: ST3  
C:Superfamily: alcohol sulfotransferase  
C:Keywords: sulfotransferase  
Alignment Scores:  
Pred. No.: 4.25e-13 Length: 325  
Score: 280.00 Matches: 82  
Percent Similarity: 46.1% Conservative: 60  
Best Local Similarity: 26.6% Mismatches: 118  
Query Match: 6.4% Indels: 48  
DB: 2 Gaps: 12  
US-10-768-158-1 (1-2419) x T07833 (1-325)  
QY 45 CCAGACCCCGGGGAGTTCGAGAGCAAGTACTTCGAGTTCCATGGC---GTGCGGCTG 101  
Db 28 ProSerGluLysGlyTrpLeuValSerGlnIleTyrGlnPheGlnGlyArgTrpHisThr 47  
QY 102 CGCCCTTCTGCGCGGGAAGATGAGAGATCGCAACTCCCGTGGCGGCCAGCGAC 161  
Db 48 GluAlaLeuLeuGlnGlyIleLeuThrCysGlnLysHisPheLysAlaLysAspSerAsp 67  
QY 162 GTGTGGATGTCACCTACCCCAAGTCCGCGCACGCTTGTCTGCAGAGGTGGTCTACTTG 221  
Db 68 IleIleLeuValThrAsnProLysSerGlyThrThrTrpLeuLysSerLeuValPheAla 87  
QY 222 -----GTGAGCCAGGGCGCTGACCCCGATGATGATCGGCTG 257  
Db 88 LeuIleAsnArgHisLysPheProValSerSerGlyAspHisPro-----LeuLeuVal 105  
QY 258 ATGAACATCGACGACGCTCCCGTCTCTGGAG-----TACCCACAGCCGCGCTG 308  
Db 106 ThrAsnProHisLeuLeuValProPheMetGluGlyValTyrTyrGluSerProAspPhe 125





probable flavonol sulfotransferase [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: E86319  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: E86319  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-346 <STO>  
 A:Cross-references: UNIPROT:Q9FZ80; UNIPARC:UPI000000C631; GB:AE005172; NID:g9795597; PI  
 C:Genetics:  
 A:Map position: 1  
 C:Superfamily: alcohol sulfotransferase

Alignment Scores:  
 Pred. No.: 5.6e-13 Length: 346  
 Score: 278.50 Matches: 83  
 Percent Similarity: 43.5% Conservative: 62  
 Best Local Similarity: 24.9% Mismatches: 119  
 Query Match: 6.3% Indels: 69  
 DB: 2 Gaps: 12

US-10-768-158-1 (1-2419) x E86319 (1-346)

QY 21 ATGGCGGAGAGAGAGCCGAG---ACCCCGACAGCCCGGGAGGAGTTGCGAG-----AGC 71  
 Db 11 ValSerGluSerAsnHisGluLeuAlaSerSerProSerGluPheGluLysAsnGln 30  
 QY 72 AAGTACTTCGAGTTCATCGCGTGGCGTCCG----- 104  
 Db 31 LysHisTyrGlnGluIleAlaThrLeuProHisLysAspGlyTrpArgProLysAsp 50  
 QY 105 -----CCCTTCGCGGGGAGAGATGAG 128  
 Db 51 ProPheValGluTyrGlyHisTrpTrpLeuGlnProLeuLeuGluGlyLeuLeuHis 70  
 QY 129 GAGATCGCCAACTTCCCGTGGCGCCGACGAGTGTGATCGTCACCTACCCCAAGTCC 188  
 Db 71 AlaGlnLysPhePheLysAlaArgProAsnAspPheValCysSerTrpProLysThr 90  
 QY 189 GGCACCACTTCGTCAGGAGTGTCTACTTGGTGAGCGAGCGGCTGACCCCGAT--- 245  
 Db 91 GlyThrThrTrpLeuLysAlaLeuThrPheAlaIleAlaAsnArgSerLysPheAspVal 110  
 QY 246 -----GAGATCGGTTGATGAACATCGACGAGAGCTCCCGTCTCGAG----- 290  
 Db 111 SerThrAsnProLeuLysArgAsnProHisGluPheValProTyrIleGluIleAsp 130  
 QY 291 TACCCA---CAGCGGCGCTGCACATCATCAGGAAGTACCTCTCCCGCTCATCAAG 347  
 Db 131 PheProPhePheProSerValAspValLeuLysAsp---GluGlyAsnThrLeuPheSer 149  
 QY 348 AGCCACCTGCGCTACCGCTTTCTGCGCTCTGACCTCCACATGGAGACTCCCAAGGTCTATC 407  
 Db 150 ThrHisIleProTyrAspLeuLeuProGluSerValValLysSerGlyCysLysIleVal 169  
 QY 408 TATATGGTTCGCAACCCCAAGGATCTGGTGTCTTATATCATGTTCCACCGCTCTCTG 467  
 Db 170 TyrIleTrpArgAspProLysAspThrPheValSerMetTrpThrPheAlaHisLysGlu 189  
 QY 468 CGGACC-----ATGAGCTACCCGAGGACCTTTCAGAATCTTCGCGGAGG 512  
 Db 190 ArgSerGlnGlnGlyProValValSerIleGluGluAlaPheAspLysTyrCysGlnGly 209

QY 513 TTTATGAATGATAAGCTGGCTACGGCTCCCTGGTTTGAGCAGCTGCAGAGTTCTGGGAG 572  
 Db 210 LeuSer-----AlaTyrGlyProTyrLeuAspHisValLeuGlyTyrTrpLys 225  
 QY 573 -----CACCGCATGACTCGAACGTGCTTTTCTCAAGTAGTGAAGACATGCATCGGAC 626  
 Db 226 AlaTyrGlnAlaAsnProAspGlnIleLeuPheLeuLysTyrGluThrMetArgAlaAsp 245  
 QY 627 CTGTTGACAGTGTGTGAGCAGCTGCCAGATTCCTGGGGGTGTCTGTGACAAAGCCCCAG 686  
 Db 246 ProLeuProTyrValLysArgLeuAlaGluPheMetGlyTyrGlyPheThrLysGluGlu 265  
 QY 687 CTGGAGCCCTGACGAGCAGCTGCCACGAGCTGGTGGACCACTGTGTCAACGCTGAGGC 746  
 Db 266 GluGluGly-----AsnValValGluLysValValLysLeuCysSerPheGluThr 282  
 QY 747 CTGCCC----- 752  
 Db 283 LeuLysAsnLeuGluAlaAsnLysGlyLysAspArgGluAspArgProAlaValTyr 302  
 QY 753 -----GTGGCGCGGAGAGAGTTCGGCTGTGGAAGGACATCTTCACCGTC 797  
 Db 303 AlaAsnSerAlaTyrPheArgLysGlyLysValGlyAspTrpGlnAsnTyrLeuThrPro 322  
 QY 798 TCCATGAATGAGAGTTTGTGCTGTGTATTAACAGAAG 836  
 Db 323 GluMetValAlaArgIleAspGlyLeuMetGluGluLys 335

RESULT 47  
 T07831  
 probable steroid sulfotransferase (EC 2.8.2.15) 1 - rape  
 C:Species: Brassica napus (rape)  
 C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 09-Jul-2004  
 C:Accession: T07831  
 R:Richard, M.; Nicolle, L.; Varin, L.  
 submitted to the EMBL Data Library, April 1997  
 A:Reference number: Z16161  
 A:Accession: T07831  
 A:Molecule type: DNA  
 A:Cross-references: UNIPROT:O82408; UNIPARC:UPI000000A1423; EMBL:AF000305; NID:g3420003;  
 C:Genetics:  
 A:Superfamily: alcohol sulfotransferase  
 C:Keywords: sulfotransferase

Alignment Scores:  
 Pred. No.: 7.82e-13 Length: 323  
 Score: 276.50 Matches: 85  
 Percent Similarity: 46.5% Conservative: 59  
 Best Local Similarity: 27.4% Mismatches: 113  
 Query Match: 6.3% Indels: 53  
 DB: 2 Gaps: 14

US-10-768-158-1 (1-2419) x T07831 (1-323)

QY 45 CCCAGCACCCCGGGGAGTTCGAGAGCAAGTACTTTCGAGTTCATGGGTG---CGGCTG 101  
 Db 27 ProSerGluLysGlyTrpLeuValSerGlnMetTyrGlnPheGluGlyIleTrpGlnThr 46  
 QY 102 CCGCCCTTCTCCCGCGGGAAGATGAGGAGATCGCAACTTCCCGGTGCGGCCCGACGAC 161  
 Db 47 GlnAlaLeuValGlnGlyIleValAsnCysGlnLysHisPheGluAlaAsnAspSerAsp 66  
 QY 162 GTGTGGATGTCACCTACCCCAAGTCCCGCACCGAGCTTGTGTGAGGAGGTGTCTACTTG 221  
 Db 67 ValIleLeuAlaThrLeuAlaLysSerGlyThrTrpTrpLeuLysAlaLeuPheAla 86  
 QY 222 GTG-----AGCCAGGCGGTGACCCCGATGAGATCGGCTTGATG 260  
 Db 87 LeuIleHisArgHisLysPheProValSerGlyLysHisPro-----LeuLeuValThr 104

```

QY 261 AACATCGACGACGAGCTCCCGGTCTCGAG-----TACCCACACCCGGGCTCGAC 311
DB 105 AsnProHisSerLeuValProTyrLeuGluGlyAspTyrCysSerSerProGluValAsn 124
QY 312 ATCATCAGGAACTGACCTCCCGGCTCATCAAGACGACCACTGCGCTACCGCTTTCTG 371
DB 125 ---PheAlaGluLeuProSerProArgLeuMetGlnThrHisLeuThrHisHisSerLeu 143
QY 372 CCTCTGACCTCCCAATGGAGACTCCCAAGGTCTATATGGTCTGCGCAACCCCAAGGAT 431
DB 144 ProValSerIleLysSerSerCysLysIleIleTyrCysArgAsnProLysAsp 163
QY 432 CTGTGGTGTCTTATATCAGTTCCACCGCTCTCTG-----CGACCATGAGCTAC 482
DB 164 MetPheValSerIleTrpHisPheGlyArgLysLeuAlaProGluLysThrAlaGluTyr 183
QY 483 -----CGAGGCACCTTTCAAGAATTCTGCGG---AGGTTTATGAATGAAGCTGGGC 533
DB 184 ProfileGluThrAlaValAlaAlaPheCysLysGlyLysPheIleGly-----199
QY 534 TACGGCTCTGTTTGGAGACGTGCGAGAGTTCTGGGACGACCGCATG-----GACTCG 587
DB 200 ---GlyProPheTrpAspHisValLeuGluTyrTrpTyrGluSerLeuLysAsnProAsn 218
QY 588 AACGTGCTTTTCTCAAGTATGAACATGATGATCGGACCTGGTGGTGGTGGAGCAG 647
DB 219 LysValLeuPheValThrTyrGluGluLeuLysLysGlnThrGluValGluValLysArg 238
QY 648 CTGGCCAGATTCCTGGGGGTGCTGTCGACAAAGGCCAGCTGGAAGCCCTGACGAGCAC 707
DB 239 IleAlaGluPheIleGly-----CysGlyPheThrAlaGluGluLysValSerGlu---255
QY 708 TGCCACCAAGCTGGTGACGAGCTGCTCAACGCTGAGGCC-----746
DB 256 -----IleValLysLeuCysSerPheGluSerLeuSerSerLeuGluValAsn 271
QY 747 -----CTGCGCGTGGGC-----CGGGGAAGAGTT 770
DB 272 ArgGlnGlyLysLeuProAsnGlyIleGluSerAsnAlaPhePheArgLysGlyLysThr 291
QY 771 GGGCTGTGAAGGACATCTTACCGCTCTCCATGAATGAGAAGTTTGACTGTGTATAAA 830
DB 292 GlyGlyTrpArgAspThrLeuSerGluSerLeuAlaAspValIleAspArgThrThrGlu 311
QY 831 CAGAAGATGGGAAGTGTGACCTCAGCTT 860
DB 312 GlnLysPheGlySerGlyLeuLysPhe 321

RESULT 48
A40216
flavonol 4'-sulfotransferase - Flaveria chloraefolia
C:Species: Flaveria chloraefolia
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
C:Accession: A40216
R:Varin, L.; Deluca, V.; Ibrahim, R.K.; Brissson, N.
Proc. Natl. Acad. Sci. U.S.A. 89, 1286-1290, 1992
A:Title: Molecular characterization of two plant flavonol sulfotransferases.
A:Reference number: A40216; MUID:92159034; PMID:1741382
A:Accession: A40216
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-320 <VAR>
A:Cross-references: UNIPARC:UPI0000175837
A:Experimental source: terminal bud
A:Note: sequence inconsistent with the nucleotide translation
A:Notes: sequence extracted from NCBI backbone (NCBIN:82216, NCBIIP:82217)
C:Superfamily: alcohol sulfotransferase

Alignment Scores:
Pred. No.: 9.3e-13 Length: 320
Score: 275.50 Matches: 80
Percent Similarity: 49.1% Conservative: 63

```

```

Best Local Similarity: 27.5% Mismatches: 111
Query Match: 6.3% Indels: 37
DB: 2 Gaps: 11

US-10-768-158-1 (1-2419) x A40216 (1-320)

QY 72 AAGTACTTCGAGTTCCATGGGTGCGGTGCGGCTTCTGCCGGGAAGATGAGGAG 131
DB 34 LysTyrGlnAspPheTrpGly-----LeuGlnAsnIleGluGlyAlaIleLeuAla 51
QY 132 ATCCGCCAACTTCCCGGTGCGGCCCGACGACGTGTGGATCGTCACCTACCCCAAGTCCGGC 191
DB 52 GlnGlnSerPheLysAlaAlaArgProAspValPheLeuCysSerTyrProLysSerGly 71
QY 192 ACCAGCTGTGTGAGGAGGTGCTACTGTGTGACGAGCGGCTGACCCCGATGAGATC 251
DB 72 ThrThrTrpLeuLysAlaLeuAlaTyrAlaIleValThrArgGluLysPheAspGluPhe 91
QY 252 GGC-----TTGATGAACATCGACGAGCAG---CTCCCGGTCTCTGGAGTACCCACAG 299
DB 92 ThrSerProLeuLeuThrAsnIleProHisAsnCysIleProTyrIleGlu-----108
QY 300 CCGGCGCTGGACATCATCAAGAACTGACCTCTCCCGC-----CTCATC 344
DB 109 -----LysAspLeuLysLysIleValAspAsnGlnAsnAsnSerCysPheThrProMet 126
QY 345 AAGAGCCACCTGCGCTTCTGCGCTTCTGCGCTGACCTCCACAATGGAGACTCCAAGTTC 404
DB 127 AlaThrHisMetProTyrHisValLeuProLysSerIleLeuAlaLeuAsnCysLysMet 146
QY 405 ATCTATATGGTTCGCAACCCCAAGATCTGTGTGTCTTATTATTCAGTTCACCGCTCT 464
DB 147 ValTyrIleTyrArgAsnIleLysAspValIleValSerPheTyrHisPheGlyArgGlu 166
QY 465 CTGCGGACCATGACGTACCGA---GGCACCTTTCAAGAAATTCGCGGAGGTTTATGAAT 521
DB 167 IleThrLysLeuProLeuGluAspAlaProPheGluGluAlaPheAspGluPheTyrHis 186
QY 522 GATAAAGCTGGGCTACGGCTCTCTGTTGACGACGTGAGAGTCTGGGAGCACCAGCATG 581
DB 187 GlyIleSerGlnPheGlyProTyrTrpAspHisLeuLeuGlyTyrTrpLysAlaSerLeu 206
QY 582 GAC-----TCGAACGTCTTTCTCAAGTATGAAGACATGCATCGATCGGAGCTGTGACG 635
DB 207 GluArgProGluValIleLeuPheLysLysTyrGluAspValLysLysAspProThrSer 226
QY 636 ATGTGTGAGCAGCTGGCCAGATTCCTGGGGGTGCTCTGT-----GACAAAG 680
DB 227 AsnValLysArgLeuAlaGluPheIleGlyTyrProPheThrPheGluGluGluLysGlu 246
QY 681 GCCAGCTGGAAGCCCTGACGAGCAGCTGC-----CACCAGCTGGTG 722
DB 247 GlyValIleGluSerIleLysLeuCysSerPheGluAsnLeuSerAsnLeuGluVal 266
QY 723 GACCAAGTGTGCAACGCTGAGGCC---CTGCCCGTG-----GGCCGG 761
DB 267 AsnLysSerGlyAsnSerLysGlyPheLeuProIleGluAsnArgLeuTyrPheArgLys 286
QY 762 GGAAGAGTTGGCTGTGGAAGACATCTTACCGCTCTCCATGAATGAGAGTTGACTTG 821
DB 287 AlaLysAspGlyAspTrpLysAsnTyrPheThrAspGluMetThrGluLysIleAspLys 306
QY 822 GTGTATAACAGAAGATGGAAAGTGTGACCTC 854
DB 307 LeuIleAspGluLysLeuSerAlaThrGlyLeu 317

RESULT 49
H96768
protein flavonol sulfotransferase E2p9.4 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: H96768
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

```





C;Accession: S69188  
R;Jacomme, C.; Roby, D.  
Plant Mol. Biol. 30, 995-1008, 1996  
A;Title: Molecular cloning of a sulfotransferase in Arabidopsis thaliana and regulation  
A;Reference number: S69188; MUID:96270377; PMID:8639757  
A;Accession: S69188  
A;Molecule type: mRNA  
A;Residues: 1-302 <IAC>  
A;Cross-references: UNIPARC:UPI0000046BEB; EMBL:Z46823; NID:g599639; PIDN:CRA86850.1; PI  
C;Superfamily: alcohol sulfotransferase  
C;Keywords: sulfotransferase

Alignment Scores:  
Pred. No.: 1.16e-11 Length: 302  
Score: 261.00 Matches: 82  
Percent Similarity: 44.7% Conservatives: 53  
Best Local Similarity: 27.2% Mismatches: 115  
Query Match: 5.9% Indels: 52  
DB: 2 Gaps: 13

US-10-768-158-1 (1-2419) x S69188 (1-302)

QY 6 GCGCAGCGCGCGCATGGCGGAGAGCGGCGGAGACCCCGACCCCGCG----- 56  
Db 12 GlyAspGluAspLeuThrGlnGluThrArgAlaLeuLeuSerSerLeuProLysGluLys 31

QY 57 GCGGAGTTCGAGAGCAAGTACTTCGAGTTCATGGCGTG---CGGCTGCCGCCCTTCTGCG 113  
Db 32 GlyTrpLeuValSerGluIleTyrGluPheGlnGlyLeuTrpHisThrGlnAlaIleLeu 51

QY 114 CCGGGGAAGATGGAGAGATCGCAACTTCCCGGTGCGGCCCGACGACGTGTGGATCGTC 173  
Db 52 GlnGlyIleLeuIleCysGlnLysArgPheGluAlaLysAspSerAspIleIleLeuVal 71

QY 174 ACCTACCCCAAGTCCGCGACCGCTTGTGCGAGGTGTCTACTTGTG----- 224  
Db 72 ThrAsnProLysSerGlyThrTrpLeuLysAlaLeuValPheAlaLeuLeuAsnArg 91

QY 225 -----AGCCAGGCGCTGACCCCGCATGATCGGTGTGATGAACATC 266  
Db 92 HisLysPheProValSerSerSerGlyAsnHisPro-----LeuLeuValThrAsnPro 109

QY 267 GAGGACGACCTCCGCTCTGGAG-----TACCACAGCGCGGCTCGACATCATC 317  
Db 110 HisLeuLeuValProPheLeuGluGlyValTyrTyrGluSerProAspPheAsp---Phe 128

QY 318 AAGGAATCAGCTCTCCCGCCTCATCAAGAGCCACTGCCCTACCGCTTCTGCCCTCT 377  
Db 129 SerSerLeuProSerProArgLeuMetAsnThrHisIleSerHisLeuSerLeuProGlu 148

QY 378 GACCTCCACATGGAGACTCCAAAGTCACTATATGGCTCGCAACCCCAAGGATCTGTG 437  
Db 149 SerValLysSerSerSerCysLysIleValTyrCysArgAsnProLysAspMetPhe 168

QY 438 GTGTCTTATATCAGTTCCACCGCTCTCTG-----CGGACCATGAGCTACCGAGCG 488  
Db 169 ValSerLeuTrpHisPheGlyLysLysLeuAlaProGluGluThrAlaAspTyr----- 186

QY 489 ACCTTTCAAGATTCTCGCGGAGTTTATGAATAGTAACTGGCTACCGCTCCTGGTTT 548  
Db 187 ProIleGluLysAlaValGluAlaPheCysGluGlyLysPheIleGlyGlyProPheTrp 206

QY 549 GAGCAGCTGAGGAGTTCTGG-----GAGCACCGCATGGACTCGAACGTGTTTTTCTC 602  
Db 207 AspHisIleLeuGluTyrTrpTyrAlaSerArgGluAsnProAsnLysValLeuPheVal 226

QY 603 AGATATGAACATGATCGGACCTGGTGACGATGGTGGAGCAGCTGGCCAGATCTCTG 662  
Db 227 ThrTyrGluGluLysLysGlnThrGluValGluMetLysArgIleAlaGluPheLeu 246

QY 663 GGGGTCTCTGTGACAAAGCCAGCTGGAAGCCCTGACCGGACGACTGCCACGCTGGTG 722  
Db 247 -----GluCysGlyPheIleGluGluGluValArgGlu-----Ile 259

QY 723 GACCACTGCTGCAACGCTGAGGCC-----CTG 749  
Db 260 ValLysLeuCysSerPheGluSerLeuSerAsnLeuGluValAsnLysGluGlyLysLeu 279

QY 750 CCGCTGGGC-----CGGGGAAGTGGCTGTGGGAAGGAC 785  
Db 280 ProAsnGlyIleGluThrLysThrPheArgLysGlyGluIleGlyTrpArgAsp 299

QY 786 ATCTTC 791  
Db 300 SerPhe 301

RESULT 53  
JC7921  
cytosolic sulfotransferase (EC 2.8.2.-) - zebra fish  
C;Species: Brachydanio rerio (zebra fish)  
C;Date: 31-Mar-2003 #sequence\_revision 31-Mar-2003 #text\_change 14-Jul-2003  
C;Accession: JC7921  
R;Sugahara, T.; Liu, C.; Pai, T.G.; Liu, M.C.  
Biochem. Biophys. Res. Commun. 300, 725-730, 2003  
A;Title: Molecular cloning, expression, and functional characterization of a novel zebr.  
A;Reference number: JC7921; MUID:22395662; PMID:12507510  
A;Accession: JC7921  
A;Molecule type: mRNA  
A;Residues: 1-296 <SUG>  
A;Cross-references: UNIPARC:UPI0000050F6E; GB:AY180110  
C;Comment: This enzyme, which belongs to the cytosolic sulfotransferase gene family tha  
s containing hydroxyl or amino group, exhibits the enzymatic activities toward a varie  
activity.  
C;Keywords: sulfotransferase

Alignment Scores:  
Pred. No.: 3.6e-11 Length: 296  
Score: 254.50 Matches: 75  
Percent Similarity: 47.2% Conservatives: 69  
Best Local Similarity: 24.6% Mismatches: 122  
Query Match: 5.8% Indels: 39  
DB: 2 Gaps: 11

US-10-768-158-1 (1-2419) x JC7921 (1-296)

QY 27 GAGAGCGAGCGCGACGCCCGGAGGATTCGAGAGCAAGTACTTCGAGTTC 86  
Db 5 LysSerArgMetGluThrAlaAlaLysMetLysAsp---GluAspLysLeuTyrArgArg 23

QY 87 CATGCG-----GTGCGGTGCGCGCTTCTGCGGGGAGATGGAGGAG 131  
Db 24 AspGlyIleLeuTyrSerThrValLeuSerProProGluThrLeuAspLysLeuLysAsp 43

QY 132 ATCGGCCAACTTCCCGTGGCGCCAGCGACGTGTGGATCGTCACCTACCCCAAGTCCGCG 191  
Db 44 Leu-----GlnAlaArgGluAspLeuIleLeuValAlaTyrProLysCysGly 60

QY 192 ACCAGCTTGTGCGAGAGGTGTC---TACTTGTGTAGCCAGGCGCTGACCCCGATGAG 248  
Db 61 PheAsnTrpMetValAlaValLeuArgLysIleAsnAlaSerThrGlyLysAspGlu 80

QY 249 ATCGGCTGTGAACATCGACGAGCAGCTCCCGGTCTCGGAGTACCCACAGCGCGGCTG 308  
Db 81 -----LysProGluArgProProLeuValGluPheLeuProThrVal 96

QY 309 ---GACATCATCAAGGAATGACCTCTCCCGCTCATCAAGAGCACCTGCCCTACCGC 365  
Db 97 GlnGluGluMetAlaGlnMetProProArgLeuLeuGlyThrHisLeuHisProAsp 116

QY 366 TTCTGCCCTCTGACCTCCAAATGGAGACTCCAAAGTCACTATATATGGCTCGCAACCC 425  
Db 117 AsnMetProAlaThrPhePheThrLysLysProLysIleLeuValValPheArgAsnPro 136

QY 426 AAGGATCTGGTGTCTTATATCAGTTCCACCCCTCTCTCGGACCATGAGTACCGA 485  
Db 137 LysAspThrLeuValSerTyrThrHisPheMetAsnLysAsnProValLeuProAsnAla 156

```

QY 486 GCACCTTTCAGAAATTCTCGCGGAGGTTTATGAATGATAAGCTGGCTACGGCTCCTCG 545
Db 157 GIUSerTTPAspLysPheSerAspPheMetThrGlyAspValSerTrpGlySerTyr 176
QY 546 TTTGAGCAGCTGAGGAGTTCTGGAGACACCGCAGTGGCTCG---AACGTGCTTTTCTC 602
Db 177 PheAspHisAlaLeuAa---TrpGluLysArgIleAspAspProAsnValMetIleVal 195
QY 603 AAGTATGAACATGATCGGACCTGCTGACGATGTTGGAGCAGCTGGCCAGATTCCTG 662
Db 196 MetTyrGluAspLeuLysGlnAanLeuProGluGlyValLysLysIleSerGluPhePhe 215
QY 663 GGGGTGCTCTGTGACAGCCAGCTGGAAGCCCTG----- 698
Db 216 SerLeuProLeuThrAspGluGlnValSerIleAlaGlyGlnSerThrPheSerAla 235
QY 699 -----ACGAGCACTGCCACACAGCTGGTGGACCACTGCTGCCAACGCTGAGGCC 746
Db 236 MetValGluAsnSerGlnLysSerHisGly-----AsnPheGlySer 249
QY 747 CTGCGCGTGGCGGGAGAGTTGGCTGTGGAAGGACATCTTCACCTCTCCATGAAT 806
Db 250 IlePhePheArgLysGlyGluValGlyAspTrpLysAsnHisPheSerGluAlaGlnSer 269
QY 807 GAGAACTTTGACTGTGTATTAACAGAAAGATGGGAAGTGTGACCTC----- 854
Db 270 LysGlnMetAspGluLeuTyrHisSerLysLeuAlaGlyThrLysLeuAlaAlaArgMet 289
QY 855 ACCTTTGACTTTTAT 869
Db 290 AsnTyrAspLeuTyr 294

RESULT 54
A96769
protein flavonol sulfoxtransferase F2p9.3 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: A96769
R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lurob, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A96769
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-338 <STO>
A:Cross-references: UNIPROT:Q9C9D0; UNIPARC:UPI00000A6FB0; GB:AE005173; NID:g7109462; P
C:Genetics:
A:Gene: F2p9.3
A:Map position: 1
C:Superfamily: alcohol sulfoxtransferase

Alignment Scores:
Pred. No.: 8.88e-11 Length: 338
Score: 249.50 Matches: 68
Percent Similarity: 46.1% Conservative: 57
Best Local Similarity: 25.1% Mismatches: 99
Query Match: 5.7% Indels: 47
DB: 2 Gaps: 10

US-10-768-158-1 (1-2419) x A96769 (1-338)

QY 138 AACTTCCCGTGGCGCCAGCGAGTGTGGATCGTCACTTACCCAGTCCGACACAGC 197
Db 66 HisPheGluAlaArgProThrAspPheLeuValCysSerTyrProLysThrGlyThrThr 85

```

```

QY 198 TTGCTGACGAGGTGGTCTACTTGTGTGAGCCAGGCGCTGACCCCGATGAGATCGGC--- 254
Db 86 TrpLeuLysAlaLeuThrTyrAlaIleValAanArgSerArgTyrAspAspAlaAan 105
QY 255 -----TTGATGAACATGACAGCAGCTCCCGGTCTCTGGAGTACCCA----- 296
Db 106 ProLeuLeuLysArgAsnProHisGluPheValProTyrValGluIleAspPheAlaPhe 125
QY 297 CAGCGCGCTGGAGCATCATCAAGAACTGCTCTCCCGCTCATCAAGAGCCACCTG 356
Db 126 TyrProThrValAspValLeuGlnAspArgLysAanPro---LeuPheSerThrHisIle 144
QY 357 CCTTACCCTTCTGCTGCTGCTGACCTCCACATGGAGACTCCCAAGTCTCATATATGGCT 416
Db 145 ProAsnGlyLeuLeuProAspSerIleValAanSerGlyCysLysMetValTyrIleTrp 164
QY 417 GCACACCAAGGATCTGGTGTCTTATTATCATGTTTC---CACCGCTCTCTG----- 467
Db 165 ArgAspProLysAspThrPheIleSerMetTrpThrPheLeuHisLysGluLysSerGln 184
QY 468 -----CGGACCATGAGCTACCGGACCTTTTCAAGAAATTTCTCCGAGGAGTTTATGAAT 521
Db 185 GluGlyGlnLeuAlaSerLeuGluAspSerPheAspMetPheCysLysGlyLeuSer--- 203
QY 522 GATAAGCTGGGTACGGCTCTCTGTTTGTGAGCAGCTGAGGAGTTCTGGGAGCACCGCATG 581
Db 204 -----ValTyrGlyProTyrLeuAspHisValLeuGlyTyrTrpLysAlaTyrGln 220
QY 582 GACTCG-----AACGTGCTTTTCTCAAGTATGAAGACATGTCATCGGACCTGGTGACG 635
Db 221 GluAsnProAspArgIleLeuPheLeuArgTyrGluThrMetArgAlaAanProLeuPro 240
QY 636 ATGTGTGAGCAGCTGCGCAGATTTCTGGGGGTGCTCC---TGTGCAAGGCCAGCCGTGNA 692
Db 241 PheValLysArgLeuAlaGluPheMetGlyTyrGlyPheThrAspGluGluGluAan 260
QY 693 GCCTTGACGGGACACTGCCACAGCTGTTGGACCACTGTCGCAACGCTGAGGCCCTGCC 752
Db 261 GlyValAlaGluLys-----ValValLysLeuCysSerPheGluThrLeuLys 276
QY 752 ----- 752
Db 277 AsnLeuGluAlaAanLysGlyAspLysGluArgGluAspArgProAlaValTyrAlaAan 296
QY 753 -----GTGGCGGGGGAAGAGTTGGCTGTGGAAGGACATCTTACCCCTCTCCATG 803
Db 297 SerAlaTyrPheArgLysGlyLysValGlyAspTrpAlaAanTyrLeuThrProGluMet 316
QY 804 AATGAGAGTTTGTGCTGTATATAACAGAAAG 836
Db 317 AlaAlaArgIleAspGlyLeuValGluGluLys 327

RESULT 55
A84523
probable steroid sulfoxtransferase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: A84523
R:Lin, X.; Kaul, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
M.; Koo, H.; Moffat, K.S.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: A84523
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333 <STO>
A:Cross-references: UNIPROT:O82330; UNIPARC:UPI000004850D; GB:AE002093; NID:g3650034; P
C:Genetics:
A:Gene: At2g14920
A:Map position: 2

```

C:Superfamily: alcohol sulfotransferase

Alignment Scores:	
Pred. No.:	4.66e-10
Score:	240.00
Percent Similarity:	44.3%
Best Local Similarity:	24.1%
Query Match:	5.5%
DB:	2
Gaps:	15
Indels:	72
Mismatches:	113
Conservative:	67
Matches:	80
Length:	333

US-10-768-158-1 (1-2419) x A84523 (1-333)

QY	27	GAGAGCGAGCGCGAGACCCCGACACCCCGGGGAGTTCGAG-----AGCAAGTAC	77
DB	20	GlutThrLysIleLeuIleSerLeuProTrpGluIleAspTyrLeuGlyAsnLysLeu	39
QY	78	TTGAGTTTCATGGCGTGGCGCTCGCGCCCTTCGCGCGGGAAGATGGAGAGATCGCC	137
DB	40	PheAsnTyrGluGlyTyrTrp-----TyrSerGluAspIleLeuGlnSerIlePro	56
QY	138	AAC-----TTCCCGGTGGCGCCACGACGCTGTGATCGTCACTACCCCAAG	185
DB	57	AsnIleHisThrGlyPheGlnProGlnThrAspIleIleLeuAlaSerPheTyrLys	76
QY	186	TCCGGCACCAAGCTTGCTGCAGAGAGTGTCTACTTGTGTGACCGAGCGGT-----	236
DB	77	SerGlyThrThrTrpLeuLysAlaLeuThrPheAlaLeuValGlnArgSerLysHisSer	96
QY	237	-----GACCCCGATGATCGCTGGCTTGATGAAC	263
DB	97	LeuGluAspHisGlnHisProLeuLeuHisHisAsnProHisGluIle---ValProAsn	115
QY	264	ATCGACGACGACGCTCCGCTCGTGAGTACCCACAGCGCGGCTGGACATCATCAGGAA	323
DB	116	LeuGluLeuAspLeu---TyrLeuLysSerLysPro-----AspLeuThrLysPhe	132
QY	324	CTG-----ACCTCTCCCGCTCATCAAGACCCACCTCGCCTACCGCTTCTTG	371
DB	133	LeuSerSerSerSerSerProArgLeuPheSerThrHisMetSerLeuAspProLeu	152
QY	372	CCCTCTGACCTCCAAATGGAGACTCCAAGTCTATATATGGCTCCGAACCCCAAGAT	431
DB	153	GlnValProLeuLysGluAsnLeuCysLysIleValTyrValCysArgAsnValLysAsp	172
QY	432	CTGGTGGTCTCTATTATCAGTTTCCACCGCTCTCTCGGG-----ACCATG	476
DB	173	ValMetValSerValTyrTyrPheArgGlnSerLysLysIleThrArgAlaGluAspTyr	192
QY	477	AGTACCGGAGGACCTTTCAAGAAATTCTCGCGGAGTTTATGAATGATAGCTGGGCTAC	536
DB	193	SerLeuGluAlaIlePheGluSerPheCys-----AsnGlyValThrLeuHis	208
QY	537	GGCTCTCGTTGTGAGCAGCTGCGAGAGTTCTGGGAGCACCGCATGGAC-----	590
DB	209	GlyProPheTrpAspHisAlaLeuSerTyrTrpArgGlySerLeuGluAspProLysHis	228
QY	591	GTGCTTTTCTCAAGTATGAACATGCATCGGACCTGGTGACGATGGTGGAGCAGCTG	650
DB	229	PheLeuPheMetCysArgTyrGluAspLeuLysAlaGluProArgThrGlnValLysArgLeu	248
QY	651	GCCAGATTCTCGGGGTGCTCTGT-----GACAAGGCCAGCTG	689
DB	249	AlaGluPheLeu-----AspCysProPheThrLysGluGluAspSerGlySerVal	266
QY	690	GAAGCCCTGACGAGCACTGCCACAGCTGTGTGGACGACGCTGTCGAACGCTGAGGCCCTG	749
DB	267	AspLysIleLeuGluLeuCys-----SerLeuSerAsnLeuArgSerVal	281
QY	750	CCGTGGGGC-----CGGGGA	764
DB	282	GluIleAsnLysThrArgThrSerSerArgValAspPheLysSerTyrPheArgLysGly	301
QY	765	AGAGTTGGGCTGTGGAAGGACATCTTCACTCGCTCCATGAATGAGAGTGTGATCTGGTG	824

```
QY 624 GACCTGGTGCAGCATGGTGGAGCAGCTGGCCAGATTCCTGGGGGTCTCTGTGCACAGGCC 683
    |||
Db 215 AspProHisGlyGlnLeuLysLysLeuAlaGluPheLeuGly-----CysProPheSer 232
    |||
QY 684 CAGCTGGAGCCCTCAGCGAGCAGCTGCACACAGCTGGTGGACCACTGTGTC-----AAC 737
    |||
Db 233 LysGluGluGluLysAsnGlySerLeuAsnLysLysLeuGluMetCysSerLeuProAsn 252
    |||
QY 738 GCTGAGGCCCTCCCGTGGCC-----758
    |||
Db 253 LeuSerSerLeuGluValAsnLysThrGlyLysSerIleAsnGlyIleGluLysAsn 272
QY 759 -----CGGGGAAGATTGGGTGTGGTGGGAAGACATCTTCACCGTCTCCATGATGAG 809
    |||
Db 273 HisPheArgLysGlyIleValGlyAspTrpLysAsnHisLeuThrProGluMetGlySer 292
QY 810 AAGTTTGACTTGCTGATATAACAGAGATG 839
    |||
Db 293 LysIleAspMetIleMetLysGluLysLeu 302

RESULT 57
H84451
probable steroid sulfotransferase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: H84451
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: H84451
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-331 <STO>
A:Cross-references: UNIPROT:Q9ZPQ6; UNIPARC:UPI00000484E3; GB:AE002093; NID:g4406766; PI
C:Genetics:
A:Gene: At2g03750
A:Map position: 2
C:Superfamily: alcohol sulfotransferase

Alignment Scores:
Pred. No.: 7e-09 Length: 331
Score: 224.50 Matches: 68
Percent Similarity: 44.6% Conservative: 53
Best Local Similarity: 25.1% Mismatches: 117
Query Match: 5.1% Indels: 33
DB: 2 Gaps: 8

US-10-768-158-1 (1-2419) x H84451 (1-331)
QY 138 AACTTCCGGTGGGCGCCAGCAGCTGGTGGTGCCTACCTACCCCAAGTCCGCGCACGAGC 197
Db 63 HisPheLysProArgAspThrAspIleLeuAlaSerLeuProLysGlyGlyThrThr 82
QY 198 TTGCTGCAGGAGGTGCTACTTGTGGAGCCAG-----GGCGTGAC 239
    |||
Db 83 TrpLeuLysSerLeuIlePheAlaValHisArgLysLysThrArgGlyThrProGln 102
QY 240 CCGCATGATAGCTGGCTGTGATGAACATCGAGCAGCAGCTCCCGTCTGGAG-----290
    |||
Db 103 ThrHisProLeuLeuGlnAsnProHisAspLeuValProPheLeuGluValGluLeu 122
QY 291 TACCCACAGCCGGCTGGACATCATCAGGAACCTGACCTCCCGCGCTCATCAAGAGC 350
    |||
Db 123 TyrAlaAsnSerGlnIleProAspLeuAlaLysTyrSerSerProMetIlePheSerThr 142
QY 351 CACCTGCCCTACCGCTTTCTGGCCTCTGACCTCCCAATGGAGACTCCAAGTCACTCAT 410
    |||
Db 143 HisMetHisLeuGlnAlaLeu---ArgGluAlaThrThrLysAlaCysLysThrValTyr 161
QY 411 ATGGCTCGCNACCCCAAGGATCTGGTGGTCTTATTATCATCTCCACCGCTCTCTG---467
```

```
Db 162 ValCysArgGlyIleLysAspThrPheValserGlyTrpHisTyrArgAsnMetLeuHis 181
    |||
QY 468 CGGACCATGAGCTACCGAGGCACCTTTCACGAATTCGCCGAGGTTTGAATGAATGAAG 527
    |||
Db 182 ArgThrLysMetAspGlnAlaThrPheGluLeuMetPheAspAlaTyr-CysArgGlyVal 201
    |||
QY 528 CTGGCTACGGCTCTCTGGTGGTGGACGCTGAGAGTTCTGGGAGCACCAGCATGGACTCG 587
    |||
Db 202 LeuLeuTyrGlyProTyrTrpGluHisValLeuSerTyrTrpLysGlySerLeuGluAla 221
    |||
QY 588 -----AACGTGCTTTTCTCAAGTATGAACACATGCATCGGACCTGCTGACGATGGTG 641
    |||
Db 222 LysGluAsnValLeuPheMetLysTyrGluGluIleLeuGluProArgValGlnVal 241
    |||
QY 642 GAGCAGCTGGCCAGATCTCTGGGGGTGCTGTGCACAAAGGCCACAGCTGGAAGCCCTGACG 701
    |||
Db 242 LysArgLeuAlaGluPheLeu-----GluCysProPheThrLysGluGluGluSer 259
    |||
QY 702 GAGCACTGCCACGCTGCTGGACAGTCTGTC-----AACGCTGAGGCCCTGCCCTG 755
    |||
Db 260 GlySerValGluGluIleLeuLysLeuCysSerLeuArgAsnLeuSerAsnLeuGluVal 279
    |||
QY 756 GGC-----CGGGGAAGA 767
    |||
Db 280 AsnLysAsnGlyThrThrArgIleGlyValAspSerGlnValPhePheArgLysGlyGlu 299
QY 768 GTTGGCTGTGGAAGGACATCTTACCGTCTCCATGAATGAGAGTTTGTGCTGTAT 827
    |||
Db 300 ValGlyAspTrpLysAsnHisLeuThrProGlnMetAlaLysThrPheAspGluIleIle 319
    |||
QY 828 AAACAGAGATCGGAAAGTGTGACCTCAGCTT 860
    |||
Db 320 AspTyrArgLeuGlyAspSerGlyLeuIlePhe 330

RESULT 58
QBEB3
BHLF1 protein - human herpesvirus 4 (strain B95-8)
C:Species: human herpesvirus 4, Epstein-Barr virus
C:Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 09-Jul-2004
C:Accession: A03742
R:Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
Mol. Biol. Med. 1, 21-45, 1983
A:Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus
A:Reference number: A93065; MUID:85035713; PMID:6092825
A:Accession: A03742
A:Molecule type: DNA
A:Residues: 1-660 <BAN>
A:Cross-references: UNIPROT:P03181; UNIPARC:UPI000000C0C0
R:Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.;
Nature 310, 207-211, 1984
A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
A:Reference number: A03794; MUID:84270667; PMID:6087149
A:Contents: annotation; protein coding region
C:Comment: The sequence contains four perfect repeats (residues 149-273, 274-398, 399-573)
C:Superfamily: human herpesvirus 4 BHLF1 protein

Alignment Scores:
Pred. No.: 3.11e-06 Length: 660
Score: 190.50 Matches: 114
Percent Similarity: 32.3% Conservative: 26
Best Local Similarity: 26.3% Mismatches: 152
Query Match: 4.3% Indels: 141
DB: 1 Gaps: 23

US-10-768-158-1 (1-2419) x QBEB3 (1-660)
QY 1079 GGAGAGCTGCAGCTTCTAAAGCGGAGACAGCT-----GCTTCGGTTGGGAATCAT 1029
    |||
Db 173 GlyAlaGlyGlnArgProSerGlyProThrGlyArgProAlaAlaProGlyAlaPro 192
    |||
QY 1028 CACACTCCCTCCGCT-----CACGGCGCTCTTCCCTTCCCGCTGTTTC 984
    |||
```



Db 193 GlyThrProAlaAlaProGlyProGlyGlyGlyAlaAlaVal-ProSerGlyAlaThrPr 212  
QY 983 ACAC----- 980  
Db 212 oHisProGluArgGlySerGlyProAlaAspProProAlaAlaAlaArgLeuProProG1 232  
QY 979 -----GCTGCTTCCAGAGTTTGTCCACAGGAA 951  
Db 232 uArgGlnGluProArgLeuProGlnAspLeuAlaAlaGlnArgCysProAlaGlyPr 252  
QY 950 TAAATGAATGCATACAGCACTTTTGGCTAGTAGACTGCTCTGGGTATTGTGAGCATGCAGG 891  
Db 252 oProProThrArgSerGlyAlaAlaAlaGlnArgThr-----HisArgArgProProG1 270  
QY 890 TTGTTGTTCTGTATTATAAATAAAGTCAACGGTAGGTCACTTTCCCATCTTCTG 831  
Db 270 yCys-----Pro-ArgSerA 275  
QY 830 TTTTATACACCAAGTCAAACTTCTCATTCATGAGAGCGGTGAAGATGCTCTTCCACAGCCC 771  
Db 275 laArgAsnProGly-----CysProArgThrTrpArg----- 286  
QY 770 AACTCTTCCCGCCACCGGAGCGCTCAGCGTTGCGACACTGGTCCACAGCTGGTG 711  
Db 287 -----ArgSerGlyAlaGlnArg--GlyHisProProProGlyAlaG 300  
QY 710 GCAGTGCTCGTCAGCGCTTCCAGCTGGGCTTGTACAGGACACCCCGAGGAATCTGGC 651  
Db 300 lyGlnArgProSerGly-----ProThrGlyGlyArgp 311  
QY 650 CAGCTGCTCCA-----CCATCGTCACAGGTCCCGATGCGATGCTTCATA 606  
Db 311 roAlaAlaProGlyAlaProGlyThrProAlaAlaProGlyProGlyGly----- 327  
QY 605 CTTGAGAAAAAGCACGTTTCAGTCCATGCGGTCTCCAGAACTCTCGCAGCGTCTCAA 546  
Db 328 -----GlyAlaAlaValProSerGlyAlaThrProHisProGluArgGlySerG 344  
QY 545 CCAGGAGCGGTACCCAGCTTATTCATTATTAACCTCCCGCAGAAAT----- 499  
Db 344 lyProAlaAspProProAla-----AlaAlaArgLeuProProGluArgG 359  
QY 498 -----CTTGAAAGGTGCTCGGTAGCTCATGGT 471  
Db 359 lnGluProArgLeuProGlnAspLeuAlaAlaGlnArgCysProAlaGlyProProP 379  
QY 470 CCGCAGAGAGCGTGGAACTGATAATAAGACACCACCATCTCTGGGGTTGC---GAGC 414  
Db 379 roThrArgSerGlyAlaAlaAlaGlnArgThrHisArgArgProProGlyCysProArgS 399  
QY 413 CATATAGATGACCT-----TGGAGTCTCCATTTGTGAGGTTCAGAGGCGAG 369  
Db 399 erAlaArgAsnProGlyCysProArgThrTrpArgArgArgSerGlyAlaGlnArgGlyH 419  
QY 368 AA-----ACCGTAGGGCAGGTGCTCTTGTATGAGCGGGGAGAGTCAGTTC 321  
Db 419 isProProGlyAlaGlyGlnArgProSerGlyProThrGlyGlyArgProAlaAlaP 439  
QY 320 CTTGATGATGTCAGGC---CCGGCTGTGGTACTCCAGGACCGGGAGCTGCTCGTCAT 264  
Db 439 ro---GlyAlaProGlyThrProAla-----AlaProGlyProGlyGlyAlaAlaV 456  
QY 263 GTTCATCAACCGCATCTCATCGGGTCAGCGGCTGGCTCACCAGTAGACCACTTCCTG 204  
Db 456 alProSerGlyAlaThrProHisProGluArgGlySerGlyProAlaAspProProAlaA 476  
QY 203 CAGCAAGCTGGTCCCGACTTGGGGTAGGTGACGATCCACAGTCTGCTGGCGGCA---- 148  
Db 476 laAlaArgLeuProProGluArgGlnGluProArgLeuProGlnAspLeuAlaAlaA 496  
QY 147 -----CCGGGAAGTTGGCGATCTCTCCATCTTCCCGCGGACAGAGGGCGGAGCCG 96  
Db 496 lnArgCysProAla-----GlyProProProThrArgSerGlyAlaAlaAlaGlnA 513

QY 95 CAGCCCATCGAACTCGAAGTACTGCTCTCGAACTCCCGGGGTCTCGGGGTCTCGGC 36  
Db 513 rgThrHis-----ArgArgProProGlyCysProArgSer---A 525  
QY 35 CTCGCTCTCGGCATCGCCGCGCTCGCGCTCGC 1  
Db 525 laArgAsnProGlyCysProArgThrTrpArgArg 536

RESULT 59

A95936  
probable alcohol sulfotransferase (EC 2.8.2.2) [imported] - Sinorhizobium meliloti (str C;Species: Sinorhizobium meliloti  
C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: A95936  
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Herna  
proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A;Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing end  
A;Reference number: A95842; MUID:21396508; PMID:11481431  
A;Accession: A95936  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-280 <KUR>  
A;Cross-references: UNIPROT:Q92VF3; UNIPARC:UPI00000CB647; GB:AL591985; PIDN:CAC49153.1  
A;Experimental sources: strain 1021, megaplasmid pSymB  
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
neault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.;  
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A;Reference number: A96039; MUID:21368234; PMID:11474104  
A;Contents: annotation  
C;Genetics:  
A;Gene: Smb21249  
A;Genome: plasmid  
C;Keywords: sulfotransferase

Alignment Scores:  
Pred. No.: 0.000449 Length: 280  
Score: 161.00 Matches: 64  
Percent Similarity: 41.0% Conservative: 46  
Best Local Similarity: 23.9% Mismatches: 98  
Query Match: 3.7% Indels: 60  
DB: 12 Gaps: 12  
US-10-768-158-1 (1-2419) x A95936 (1-280)

QY 156 ACGCAGCTGTGGATCGTCACTACCCCAAGTCCGCGCACCGAGTTGTCGAGGAGTGGTC 215  
Db 25 AlaAspSerPheLeuIleSerTyrProlySerGlyArgThrTrpPheArgTyrValLeu 44  
QY 216 -----TACTTGTGAGCCAGCGGCGCTGACCCCGATGAGATCGGTTGATGAACTCGAC 269  
Db 45 SerHisTyrLeuAlaThrIleAlaArgValProGluThrIleAspLeuHisAsnMetPhe 64  
QY 270 GAGCAGCTCCCGGTCTGGAGTACCCACAGCGGCGCTGACATCATCAAGGAACGTGACC 329  
Db 65 SerIleValProAsnPheAsp-----LeuAspProValArgGlyMetPro 79  
QY 330 TCT-----CCCGGCTCATCAAGACCCACTTGC 359  
Db 80 GlyTyrArgPheArgGluAlaIleProThrIleLeuValSerHisLeuAsp 99  
QY 360 TACCGC-----TTTCTGCCCTCTGACCTCCCAATGGAGACTCCAGGTCTATCAT 410  
Db 100 TyrArgAlaSerLeuPheLeu-----ArgArgProValIleMet 112  
QY 411 ATGGCTCCGAACCCCAAGATCTGGTGTCTTATTATCATAGTTCACCGCTCTCGCG 470  
Db 113 MetValArgAspProArgAspValIleValSerAlaTyr---PheHisAlaThrArgHis 131





Db 234 -----GlyGlyProLeuProGlnProProProProProPro----- 244  
 QY 405 TGACCTGGAGTCTCCATTGTGGAGTCAGAGGGCAGAAAGCGGTAGGCGAGGTGGCTCT 346  
 Db 245 -----GlyArgSerArgProAlaAlaAla----- 254  
 QY 345 TGATAGCGGGGAGAGGTGACGTCTTGTATGATGATCCAGGCCGGCTGTGGTACTCCA 286  
 Db 255 -----ProProProAlaGluGlyThr--- 261  
 QY 285 GGACCGGAGCTGCTCGTCGATGTTTCATCAAGCCGATCTTCGCGGGTCAGGCCCTGGC 226  
 Db 262 -----AlaValThrThrThrAlaSerProTrpLeuAspGluProAla 278  
 QY 225 TCACCAAGTAGACCACTCTGTCACCAAGCTGTCGCGAGCTGGGTAGGTGACGATCC 166  
 Db 279 AlaAlaArgLeuAspProAlaAlaAlaTrp-----ArgPro 291  
 QY 165 ACACGCTCGTGGCGCCGACCGGAGTTGGCGATCTCTCCATCTCCGCGGCGAAGG 106  
 Db 292 GluProArgLeuLeuGlnProGlnLeuGlnLeuGlnLeuHisHisArgArgArg 311  
 QY 105 GGGCAGCGCGACGCATCGAACTCGAAGTACTTGTCTCGAACTCCCGCGGGTGTGG 46  
 Db 312 AlaArgArgProArg-----ProArgGluGlyArgGlyArgPro----- 326  
 QY 45 GGGTCTCGGCTCTCTCCGATCGCGCATCGCGCGCGCTGCGCGTGGC 1  
 Db 327 ---ArgArgGlyArgGlyAlaProLeuGlnArgGlnProProArgArg 340  
 RESULT 63  
 S41067  
 collagen alpha 1(III) chain - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text\_change 09-Jul-2004  
 C:Accession: S41067; A29905; S31924  
 R:Glumoff, V.; Maekelae, J.K.; Vuorio, E.  
 Biochim. Biophys. Acta 1217, 41-48, 1994  
 A:Title: Cloning of cDNA for rat pro alpha-1(III) collagen mRNA. Different expression pa  
 A:Reference number: S41067; MUID:94114571; PMID:8286415  
 A:Accession: S41067  
 A>Status: Preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-636 <GLU>  
 A:Cross-references: UNIPROT:P13941; UNIPARC:UPI0000126D1A; EMBL:X70369; NID:957915; PIDN  
 R:Frankel, F.R.; Hsu, C.Y.J.; Myers, J.C.; Lin, E.; Lyttle, C.R.; Komm, B.; Mohn, K.  
 DNA 7, 347-354, 1988  
 A:Title: Regulation of alpha-2 (I), alpha-1 (III), and alpha-2 (V) collagen mRNAs by estr  
 A:Reference number: A29905; MUID:88296083; PMID:2456504  
 A:Accession: A29905  
 A:Molecule type: mRNA  
 A:Residues: 308-482 <FRA>  
 A:Cross-references: UNIPARC:UPI0000170BC0; GB:M21354; NID:G203500; PIDN:AAA40942.1; PID:  
 R:Glumoff, V.; Maekelae, J.K.; Vuorio, E.  
 submitted to the EMBL Data Library, February 1993  
 A:Reference number: S31924  
 A:Accession: S31924  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 2-636 <GL2>  
 A:Cross-references: UNIPARC:UPI000017737E; EMBL:X70369  
 C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;  
 F:408-636/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Alignment Scores:  
 Pred. No.: 0.00141 Length: 636  
 Score: 155.50 Matches: 86  
 Percent Similarity: 29.5% Conservative: 8  
 Best Local Similarity: 27.0% Mismatches: 114  
 Query Match: 3.5% Indels: 111  
 DB: 2 Gaps: 15

US-10-768-158-1 (1-2419) x S41067 (1-636)  
 QY 762 CCCGCCCCAGGGCAGGGCCTCAGCGTTGTCAGACTGTCACACAGCTGGTGGCAGTGCCT 703  
 Db 65 ProGlyProSerGlyAlaProGlyLysAspGlyProProGlyProAlaGlyAsnSerGly 84  
 QY 702 CCCTCAGCGCTTCCA-----GCTGGG-----CCTTGTCCACAGG 670  
 Db 85 SerProGlyAsnProGlyValAlaGlyProLysGlyAspAlaGlyGlnProGlyGluLys 104  
 QY 669 ACACCCCGAGGAATCTGGCCAGCTGCTCCACCATCTGTCACCAAGT----- 625  
 Db 105 GlyProProGly-----AlaGlnGlyProProGlySerProGlyProLeuGlyIleAla 122  
 QY 625 ----- 625  
 Db 123 GlyLeuThrGlyAlaArgGlyLeuAlaGlyProProGlyMetProGlyProArgGlySer 142  
 QY 624 -----CCCGATGCATGCTTTCATACTTGA 601  
 Db 143 ProGlyProGlnGlyIleLysGlyGluSerGlyLysProGlyAlaSerGlyHisAsnGly 162  
 QY 600 GAAAGAACACGTTCCAGTCCATGCGGTGCTCCAGAACTCTCTGCACGCTGCTCAAAACCAAG 541  
 Db 163 GluArgGlyProProGlyProGlnGlyLeuProGlyGlnProGlyThrAlaGlyGluPro 182  
 QY 540 AGCGGT-----AGCCAGCTTATCATTCATAAACCTCCGCGC 505  
 Db 183 GlyArgAspGlyAsnProGlySerAspGlyGlnProGlyArgAspGlySerProGlyGly 202  
 QY 504 AGAATCTTTGAAAGGTGCTCGTAGCTCATGTCGCGCAGAGAGCGGTGGAATGATAAT 445  
 Db 203 LysGlyAspArgGlyGluAsnGlySer-----ProGlyAlaProGlyAlaProGlyHis 220  
 QY 444 AAGACACCAACCATGCTCTGGGTGTCGAGCCATATAGATGACCTTGGAGTCTCCATTGT 385  
 Db 221 ProGlyProProGlyProValGlyProSerGlyLys----- 232  
 QY 384 GGAGTCCAGAGGCGAAGAGCGGTAGGCGAGTGGCTCTTATGAGCGGGGAGAGGTCA 325  
 Db 233 AsnGlyAspArgGlyGluThrGlyProAlaGly----- 243  
 QY 324 GTTCCTTGTATGATGTCACAGCCCG-----GCTGTGGTACTCTCCAGGACCGGAGTGT 271  
 Db 244 ---ProSerGlyAlaProGlyProAlaGlyAlaArgGlyAlaProGlyProGlnGlyPro 262  
 QY 270 GGTGATGTTTCATCAAGCGATCTCATCGGGTACAGCG----- 232  
 Db 263 ArgGlyAspLysGlyGluThrGlyGluArgGlySerAsnGlyIleLysGlyHisArgGly 282  
 QY 231 ---CCTGGCTACCAAGTAGACCA-----CCTCTGCGAAGCTGGTCCGGACTTGGGT 178  
 Db 283 PheProGlyAsnProGlyProProGlySerProGlyAla-AlaGlyHisGlnGlyAlaVa 302  
 QY 177 AGGTGACGATCCACACGTCGCTGGCGCGCAGCGGAAGTGGCGATCTCTCCATCTTCC 118  
 Db 302 IGly---SerProGlyProAlaGlyProArgGlyProValGlyPro-----HisGlyPr 319  
 QY 117 CGCGCAGAGGCGCGCAGCGCACGCCCATGGAATCGAAGTACTGCTCTCGAACTCC 58  
 Db 319 oProGly-LysAspGlySer-----SerGlyHisPr 329  
 QY 57 CGCGGGTCTGGGGTCTCGGCTCGCTCTCCGCCATCGCGCGCGCTCGCGC 5  
 Db 329 roGlyProIleGly-----ProProGlyPro 337

## RESULT 64

PIHUB6

salivary proline-rich phosphoprotein precursor PRB1 (large allele) [validated] - human  
 N:Contains: peptide IB-1; peptide P-E (peptide IB-9); peptide P-F; peptide P-H  
 C:Species: Homo sapiens (man)  
 C:Date: 04-Dec-1986 #sequence revision 12-Apr-1996 #text change 05-Oct-2004  
 C:Accession: B40750; C40750; A40750; C25372; S02128; A03293; A90502; A91974; A0



```
Db 236 ProProGly---LyProGlnGlyProProGlnGln-----GlyGlyAsnArgPro 251
Qy 369 GAAAGCGTAGGCGAGGTGGCTTTGATGAGCGCGGGAGAGTCAAGTCTTCCTTGATGATGT 310
Db 252 GlnGlyProProPro----- 256
Qy 309 CCAGCCCGGCTGTGGTACTCCAGGACCGGGAGCTGCTCGCATGTTTCATCAAGCCGA 250
Db 257 ProGlyLyProGlnGlyProProProGlnGlyAspLySerArgSerProGlnSer--- 275
Qy 249 TCTCATCGGGTCAGCGCCCTCGCTCACCACAGTACACCACTCTCCGCGAGCTGGTGC 190
Db 276 -----ProProGlyLyProGlnGlyProProProGln----- 286
Qy 189 CGGACTTGGGTAGTGACGATCCACGTCCTCGGCGCACCGGGAAGTTGGCGATCT 130
Db 287 -----GlyGlyAsnGlnProGlnGlyProProProProGlyLysProGlnGly 303
Qy 129 CTTCATCTTCCCGCGCGAGAGG----- 303
Db 304 ProProProGlnGlyGlyAsnLysProGlnGlyProProProProGlyLysProGlnGly 323
Qy 102 GCAGCGCGACGATGGAAGTCTGCTCTCGAAGTCTTCTCGAAGTCTCCCGCGGGTCTGGGGG 43
Db 324 ProProAlaGlnGlyGlySerLysSerGlnSerAlaArgAlaProProGly----- 340
Qy 42 TCTCGGCTCTCGTCTCCGCCA 22
Db 341 ---LyProGlnGlyProPro 346
RESULT 65
B40505
Hypothetical protein - suid herpesvirus 1 (strain Indiana-Funkhuser or Becker)
C:Species: suid herpesvirus 1
C:Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 09-Jul-2004
C:Accession: B40505
R:Cheung, A.K.
J. Virol. 65, 5260-5271, 1991
A:Title: Cloning of the latency gene and the early protein 0 gene of pseudorabies virus.
A:Reference number: A40505; MUID:91374576; PMID:1654441
A:Accession: B40505
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1958 <CHB>
A:Cross-references: UNIPROT:Q69340; UNIPARC:UPI00000F2308; GB:M57505; NID:g334066; PIDN:
C:Superfamily: pseudorabies virus 1 nuclear antigen
Alignment Scores:
Pred. No.: 0.00214 Length: 1958
Score: 154.50 Matches: 83
Percent Similarity: 31.5% Conservative: 18
Best Local Similarity: 25.9% Mismatches: 112
Query Match: 3.5% Indels: 108
DB: 2 Gaps: 15
US-10-768-158-1 (1-2419) x B40505 (1-1958)
Qy 849 CACACTTCCCATCTTCTGTTTATACACCAAGTCAAACTTCTCAT----- 805
Db 317 HisGlnAspProProGlyProProThrProSerThrSerHisHisHisHisGln 336
Qy 804 -----TCATGGAGACGGTGAAGATGCTCTCCACAGCCCACTCTTCCCGCGC 757
Db 337 GlyProProThrSerProArg-----ProSerThrSerHisGlnAspPro 352
Qy 756 CCACGGCGAGGGCTCAGCGTTGCAGCACT-----GTTCCACCAAGCT 715
Db 353 ProGlyGlyGlyProProSerAlaGluThrHisHisHisHisHisGlnAspProProGly 372
Qy 714 GTGGGAGTGTCTCGTCAGGGCTTCCAGCTGGCGCTTGCACAGGACACCCCGAAGTCC 655
Db 373 GlyGly-----ProProSerThrSerSerHisHisHisHisGlnAspProProGlyGly 390
```

```
Qy 654 TGGCAGCTGCTCCACCATGCTCACCAGTCCCGATGTCATGCTCTTCATCTTGAGAAAAA 595
Db 391 GlyProProSerProProProArgProSerThrSerSerSer-----SerSer 406
Qy 594 GCAGTTTCGAGTCCATCGGTGCTCCAGAACTCTCGCACGTCGTCTCAACACGAGGACCGT 535
Db 407 HisGlnGlyProProSerThrArgProProProGlnArgProProProArgTrpPro 426
Qy 534 AGCCAGCTTATCATTTATAAACCTCCGGCAGATTCTTGAAGGTGCTCGGTAGCTCA 475
Db 427 ProPro-----SerProGlnLysLysSerGluThrArgAlaGlySerGlu 441
Qy 474 TGGTCCGAGAGAGG-----GGTGAACATATAAGACACACACAGATCTTGGGGT 421
Db 442 AsnThrAlaGlnThrLeuPheSerHisSerGluAsnLysLeuPheSerHisProMetGly 461
Qy 420 TCGAGCCATATAGATGACCTTGGAGTCTCATTTGTGGAGTCCAGAGGCGAGAACCGGT 361
Db 462 -----GluGlyGlyGluGlyAspArgGly 469
Qy 360 AGGCGAGGTGGCTCTTGATGAGGGCGGGAGAGGTCA-----GTTCTTGA 316
Db 470 ThrAlaGlyGlyGlu-----GlyAspArgAspAspProArgProProSerProPro 486
Qy 315 TGATGTCCAGCGCGGCTGTCGGTACTCCAGGACCGGAGCTGCTCGTCGATGTTTCATCA 256
Db 487 ProArgProProProProProProProProProPro----- 498
Qy 255 AGCCGATCTCATCGGGGTCCAGCGGCTCGGTCCACCAAGTAGACCACTCTTCACAGAACG 196
Db 499 -----ProProProProGlnProProProAlaGly--- 509
Qy 195 TGGTCCCGACTTGGGTAGTGACGATCCACACGTCGCTGGCGCGCACCGGAGTTGG 136
Db 510 -----GlySerAlaArgArgArgArg-----GlyGlyGly 520
Qy 135 CGATCTCTCCATCTTCCCGGCGGAGAGGGCGGCGGCGCACGC-----CATGGA 85
Db 521 GlyGlyProProGlyArgGlyArgArgArgGlyGlyLysArgArgAlaGluGly 540
Qy 84 ACTCGAAGTACTTGTCTCTC----- 66
Db 541 Thr-GluAlaAlaAlaAlaAspAlaGluGluGluAspGlyAspGlyAspGluAspGlu 560
Qy 65 -----GAATCTCCCGCGGCTGCTGGGG 43
Db 560 uAspArgAlaGluAspGluArgGluAspGlyGlyGluGlyProArgGlyAlaGlyGly 580
Qy 42 T 42
Db 580 Y 580
```

## RESULT 66

WMBE38

infected cell protein ICP34.5 - human herpesvirus 1 (strain F)

C:Species: human herpesvirus 1

C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 09-Jul-2004

C:Accession: A27768

R:Chou, J.; Roizman, B.

J. Virol. 57, 629-637, 1986

A:Title: The terminal a sequence of the herpes simplex virus genome contains the promoter:

A:Reference number: A27768; MUID:86115412; PMID:3003394

A:Accession: A27768

A:Molecule type: DNA

A:Residues: 1-358 &lt;CHO&gt;

A:Cross-references: UNIPROT:P08353; UNIPARC:UPI0000170E8C; GB:M12240; NID:g330122; PIDN:

C:Comment: This protein accumulates late in infection predominantly in the cytoplasm of

C:Genetics:

A:Gene: ICP34.5

C:Superfamily: herpesvirus infected cell protein ICP34.5

C:Keywords: tandem repeat

F:175-204/Region: 3-residue repeats (A-T-P)

Alignment Scores:									
Pred. No.:	0.00176	Length:	358						
Score:	153.50	Matches:	83						
Percent Similarity:	32.9%	Conservative:	18						
Best Local Similarity:	27.0%	Mismatches:	120						
Query Match:	3.5%	Indels:	86						
DB:	1	Gaps:	16						
US-10-768-158-1 (1-2419) x WMBE38 (1-358)									
QY	772	CCAACCTCTTCCCGGCCACGGCGCTCAGCGTTGCAGCACTGGTCCACCACTGG	713						
DB	25	ProArgArgProArgProGlyProThrGlyAlaVal-ProThrAlaGlnProGlnVa	44						
QY	712	TGGCAGTGTCTCGTTCAGGCTTCCAGCTGGCGCTTGTTCACAGACACACCCAGGAATCTG	653						
DB	44	lThrSerThrProAenSerGluProAla-----ValArgSerAlaProAlaAlaA	61						
QY	652	GCCAGTGTCTCCACCATCTGCACAGGTCCCGATGTCATGT	613						
DB	61	aProProProProAlaSerGlyProProProSerCysSerLeuLeuLeuArgGlnTr	81						
QY	612	CTTTCATCTTGCAGAAAAGCAGCTTCGAGT-----	583						
DB	81	pLeuHisValProGluSerAlaSerAspAspAspAspTrrProAspSerPr	101						
QY	582	----CCATGCGGTGTCTCCAGAACT---CCTGCAGGTGCTCAACACGAGGCGCTAGCCC	530						
DB	101	oProProGluProAlaProGluAlaGlyProProAlaAlaAlaProArgProArgSerPr	121						
QY	529	ACGTATCATTCATAAACCTCCGGCAGAAATCTTGAAGGTGCCTCGGTAGCTCATGGTC	470						
DB	121	o-----	125						
QY	469	CCGAGAGACGGTGAACGTGATAATAAGACACACAGATCTTGGGGTTTCGAGCCATA	410						
DB	125	aGlyProGlyGlyAlaAsnProSerHisProProSer-----ArgProPh	141						
QY	409	TAGATGACCTTGGAGTCTC-----CATGTGGAGGTCA	377						
DB	141	eArgLeuProProArgLeuAlaLeuArgLeuArgValThrAlaGluHisLeuAlaArgLe	161						
QY	376	GAGGCGAGAAACGGTAGGCGAGGTGCTTTGATGAGCGGGAGAGGTCAAGTCTCTTG	317						
DB	161	uArgArgAspAla---ArgAlaGlyGlyGly-----AlaGlyAlaProAlaThrProAl	178						
QY	316	ATGATGTCTCAGGCGCGGTGTG---GGTACTCCAGACCGGAGCTGCTCGTCGATGTC	260						
DB	178	aThrProAlaThrProAlaThrProAlaThrProAlaThrProAlaThrProAlaThrPr	198						
QY	259	ATCAAGCCGATCTCATCGGGGTGAGCGCCCTGGCTCACCA-----AGTAGACC	212						
DB	198	oAlaThr-----ProAlaThrProAlaArgValArgSerArgPr	211						
QY	211	ACCTCTCGCAGCAAGCTGGTCCGCACTTG-----GGTAGGTGACGATCCACACG	161						
DB	211	oThrValArgValArgHisLeuValValTrpAlaSerAlaAlaArgLeuArgAlaAlaA	231						
QY	160	TCGCTGGGCGCACCGGGA-----AGTTGGCGATCTCTCTCC	125						
DB	231	aArgGlyProAlaSerGlyProThrGlyLeuGlySerGlyAlaGlyTrpArgArgProAr	251						
QY	124	ATCTTCCCGCGCAGA-----AGGGCGGCGAGCGCGCAGCCATGG	86						
DB	251	gArgSerSerGlyArgAlaTrpGlyProGlyProCysArgAlaLeuProAlaGlyProAl	271						
QY	85	AACTCGAAGTACTGCTCTCGAACTCCCGGGGTCTGGGGGTCTCGGCCT---CGCTC	29						
DB	271	aArgArgThrArgSerAsnValThrProGluAlaAlaTrpSerSerAlaGluLeuArgTh	291						
QY	28	TCGCCATGCGCCGCGCT	10						

Db 291 rLysProLeuSerArgArg 297

RESULT 67

EDBE11

immediate-early protein IE110 - human herpesvirus 1 (strain 17)

C:Species: human herpesvirus 1

C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 09-Jul-2004

C:Accession: A29152

R:Perry, L.J.; Rixon, F.J.; Everett, R.D.; Frame, M.C.; McGeoch, D.J.

J. Gen. Virol. 67, 2365-2380, 1986

A:Title: Characterization of the IE110 gene of herpes simplex virus type 1.

A:Reference number: A29152; MUID:87059760; PMID:3023529

A:Accession: A29152

A:Molecule type: DNA

A:Residues: 1-775 <PER>

A:Cross-references: UNIPROT:P08393; UNIPARC:UPI000012D178; GB:X04614; NID:g59832; PIDN:

C:Genetics:

A:Introns: 19/3; 242/1

C:Superfamily: herpesvirus immediate-early protein IE110; RING finger homology

C:Keywords: DNA binding; early protein; transcription regulation; zinc finger

F:112-162/Domain: RING finger homology <RING>

F:116-156/Region: zinc finger C3HC4 motif

Alignment Scores:

Pred. No.: 0.00208 Length: 775

Score: 153.50 Matches: 68

Percent Similarity: 31.9% Conservative: 12

Best Local Similarity: 27.1% Mismatches: 82

Query Match: 3.5% Indels: 89

DB: 1 Gaps: 11

US-10-768-158-1 (1-2419) x EDBE11 (1-775)

QY	756	CCACGGGAGGGCCCTCAGCGTTGCAGCACTGTGTCCACAGCTGTGGGAGTGTCCGTCA	697
DB	253	ProArgAlaProProArgArgGlyAlaAlaAlaProProValThrGlyGlyAlaSerHis	272
QY	696	GGGCTTCAGCTGGGCGCTTGTTCACAGGACACCCCGCAGGAATCTGGCCAGCTCTCCACCA	637
DB	273	AlaAlaProGlnProAlaAlaAlaArgThr-----AlaProPro	285
QY	636	TGCTCACCA---GGTCCCGATCATGTCTTCATCTTACAGAAAAAGCAGCTTCGAGTCCA	580
DB	286	SerAlaProIleGlyProHisGlySerSerAsnThrThrThrAsnSerSerGly	305
QY	579	TGCGTGTCTCCAGAACTCCTGCAGCTCTCAAAACAGAGCGGTAGCCAGCTTATCAT	520
DB	306	GlyGlyGlySerArg-----GlnSerArgAlaAlaAlaPro-----	317
QY	519	TCATAAACCTCCGGCAGAAATCTTGAAGGTGCCTCGGTAGCTCATGTCCGACAGAGAC	460
DB	318	-----ArgGlyAlaSerGlyProSer	324
QY	459	GGTGAACCTGATATAAGACACACCACCATCTCTGGGGTTGGAGCCATATAGTACCT	400
DB	325	GlyGlyVal-----	327
QY	399	TGGAGTCTCCATTGTGGAGGTGACAGAGCGCAGAAAGCGGTAGGCGAGGTGCTTTGATCA	340
DB	328	---GlyValGlyValGlyValGluAlaGluAlaGlyArgProArgGlyArg-----	344
QY	339	GGCGGGGAGAGGTGAGTTCTCTTGATGATGTCCAGCGCCGGCTGTGGGTACTCTCCAGAGCG	280
DB	345	-----ThrGlyProLeuValAsnArgProAlaPro	354
QY	279	GGAGTGTCTCTCA-----TGTTCATCAAGCCGATCTCATCGGGGTGAGCGC	232
DB	355	LeuAlaAsnAsnArgAspProIleValIleSerAspSer-----Pro	368
QY	231	CCTGGCTCACCAAGTAGACCACTCTCGCAACAAGCTGGTCCGCGACTTGGGGTAGGTA	172
DB	369	ProAlaSerProHisArgProProAlaAla-----	378

QY 171 CGATCCACAGTCGTGGCGGCACCGGAGT-----TGGCGATCTCTCCACTTCC 118  
Db 379 -----ProMetProGlySerAlaProArgProGlyProProAlaSer 392  
QY 117 CGCGCAGAGGCGCGCAGCCAGCCATCGAAGTCTGCTCTCGAACTCC 58  
Db 393 AlaAlaAlaSerGlyProAlaArgProArgAlaAlaValAlaProCysValArgAlaPro 412  
QY 57 CGCGGTCTGGGCTCTGGCGCTCGCTCTCG 25  
Db 413 ProPro-----GlyProGlyProArgAlaPro 421  
RESULT 68  
S21626  
collagen alpha 1(I) chain precursor - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 13-Jan-1995 #sequence revision 25-Apr-1997 #text change 09-Jul-2004  
C:Accession: S57243; S16374; A23982; I49559; I49557; S39789; I48300; S21626  
R:Li, S.W.; Khillan, J.; Prockop, D.J.  
Matrix Biol. 14, 593-595, 1994  
A:Title: The complete cDNA coding sequence for the mouse pro-alpha-1(I) chain of type I  
A:Reference number: S57243  
A:Accession: S57243  
A:Molecule type: mRNA  
A:Residues: 1-1453 <LIS>  
A:Cross-references: UNIPROT:P11087; UNIPARC:UPI0000027558; EMBL:U08020; NID:g470673; PID  
R:Metaaranta, M.; Toman, D.; de Crombrughe, B.; Vuorio, E.  
Biochim. Biophys. Acta 1089, 241-243, 1991  
A:Title: Specific hybridization probes for mouse type I, II, III and IX collagen mRNAs.  
A:Reference number: S16176; MUID:91274355; PMID:2054384  
A:Accession: S16374  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1442-1453 <MET>  
A:Cross-references: UNIPARC:UPI0000000691; EMBL:X57981; NID:g50484; PIDN:CAA1046.1; PID  
R:French, B.T.; Lee, W.H.; Maul, G.G.  
Gene 39, 311-312, 1985  
A:Title: Nucleotide sequence of a cDNA clone for mouse proalpha1(I) collagen protein.  
A:Reference number: A23982; MUID:86137403; PMID:3841523  
A:Accession: A23982  
A:Molecule type: mRNA  
A:Residues: 518-1128 <PRE>  
A:Cross-references: UNIPARC:UPI000016CC46; GB:M14423; NID:g192261; PIDN:AAA37333.1; PID:  
R:Monson, J.M.; Friedman, J.; McCarthy, B.J.  
Mol. Cell. Biol. 2, 1362-1371, 1982  
A:Title: DNA sequence analysis of a mouse pro-alpha-1(I) procollagen gene: Evidence for  
A:Reference number: I49559; MUID:83141374; PMID:6298597  
A:Accession: I49559  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 735-1130 <RES>  
A:Cross-references: UNIPARC:UPI000016CC47; GB:M17491; NID:g192263; PIDN:AAA37334.1; PID:  
R:Harbers, K.; Kuehn, M.; Delius, H.; Jaenisch, R.  
Proc. Natl. Acad. Sci. U.S.A. 81, 1504-1508, 1984  
A:Title: Insertion of retrovirus into the first intron of alpha1(I) collagen gene leads  
A:Reference number: I49557; MUID:84170331; PMID:6324198  
A:Accession: I49557  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-25 <RE2>  
A:Cross-references: UNIPARC:UPI000005B79; GB:K01688; NID:g192246; PIDN:AAA37330.1; PID:  
R:Fenton, S.P.; Llannde, S.R.; Hannagan, M.; Stacey, A.; Jaenisch, R.; Bateman, J.F.  
Biochim. Biophys. Acta 1216, 469-474, 1993  
A:Title: Genomic sequence of mouse COL1A1 encoding the collagen propeptides.  
A:Reference number: S39789; MUID:94092741; PMID:8268229  
A:Accession: S39789  
A:Molecule type: DNA  
A:Residues: 1-80, 'E', 82-105, 'D', 107-185, 1031-1201, 'G', 1203-1218, 'E', 1220-1221, 'T', 1223-1  
A:Cross-references: UNIPARC:UPI000005B65; UNIPARC:UPI000017738A  
R:Rhodes, K.; Rippe, R.A.; Umezawa, A.; Nehls, M.; Brenner, D.A.; Breindl, M.  
Mol. Cell. Biol. 14, 5950-5960, 1994  
A:Title: DNA methylation represses the murine alpha 1(I) collagen promoter by an indirect  
A:Reference number: I48300; MUID:94344105; PMID:8065328

A:Accession: I48300  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>  
A:Cross-references: UNIPARC:UPI000017738B; EMBL:X54876; NID:g50486; PIDN:CAA38657.1; PID  
C:Genetics:  
A:Gene: COL1A1  
A:Introns: 770/3; 788/3; 806/3; 842/3; 860/3; 932/3; 968/3; 1004/3; 1023/3; 1058/3; 1058/3  
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;  
C:Keywords: coiled coil; extracellular matrix; glycoprotein; heterotrimer; triple helix  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-151/Domain: amino-terminal propeptide #status predicted <PRO>  
F:30-89/Domain: von Willebrand factor type C repeat homology <VMC>  
F:152-1453/Product: collagen alpha 1(I) chain #status predicted <MAT>  
F:1224-1453/Domain: fibrillar collagen carboxyl-terminal homology <PCC>  
Alignment Scores:  
Pred. No.: 0.00239 Length: 1453  
Score: 153.50 Matches: 84  
Percent Similarity: 28.9% Conservative: 8  
Best Local Similarity: 26.4% Mismatches: 119  
Query Match: 3.5% Indels: 107  
DB: 2 Gaps: 18  
US-10-768-158-1 (1-2419) x S21626 (1-1453)  
QY 783 CTTCCACAGCCCAACTCTTCCCGGCCCA-----CGGCAGGCGCTCCGGTTGCAGC 730  
Db 859 ProProGlyAlaThrGlyPheProGlyAlaAlaGlyArgGlyProProGlyProSer 878  
QY 729 -----ACTGGTCCACAGCT-----GGTGGCAGTCTCTCCG 700  
Db 879 GlyAsnAlaGlyProProGlyProProGlyProValGlyLysGluGlyLysGlyPro 898  
QY 699 TCAGGGCTCCAGCTGGCGCTTGTCCACAGACA-----CCCCCA 661  
Db 899 ArgGly---GluThrGlyProAlaGlyArgProGlyGluValGlyProProGlyPro 917  
QY 660 GGAATCTGGCAGCTGTCCACCATCGTCACAGGTCCGATGCGATGCTTCATCTTGA 601  
Db 918 Gly-----ProAlaGlyGluGlySerProGlyAlaAsp----- 929  
QY 600 GAAAAAGCACGTTCCAGTCCATCGCGTCTCCAGAACTCTCCACGTGCTCAAAACAGG 541  
Db 930 -----GlyProAlaGlySerProGlyThrProGlyProGlnGlyLeuAla 944  
QY 540 AGCGGTAGCCAGCTTATCATTAACCTCCGCGAGAAATCTTTGAAGGTGCCTC--- 484  
Db 945 GlyGlnArgGlyValValGly-----LeuProGlyGlnArgGlyGluArgGlyPhePro 962  
QY 483 -----GGTAGCTCATGG-----TCCGCAGAGCGGTGGAACTGATAAT 445  
Db 963 GlyLeuProGlyProSerGlyGluProGlyLysGlnGlyProSerGlySerGlyGlu 982  
QY 444 AAGACACACACAGATCTTGGGGTTGCGAGCCATATAGATCACCTTGGAGTCTCCATTGT 385  
Db 983 ArgGlyProProGlyProMetGly-----ProProGlyLeu----- 994  
QY 384 GGAGGTACAGGGCGGAGAAAGCGGTAGGCGAGGTGGCTCTTGTATAGGCGGGAGAGGTCA 325  
Db 995 AlaGlyProProGlyGluSerGlyArgGluGlySer----- 1006  
QY 324 GTTCTTGTATGATGTCAGGCCCGCTGTGGGTACTCCAGGACCGGAGCTGCTGTCGA 265  
Db 1007 -----ProGly---AlaGluGlySerProGlyArgAspGlyAlaProGly 1020  
QY 264 TGTTCATCAAGCCGATCTCATCGGGTCCAG-----CGCCTCGCTCCACA 220  
Db 1021 AlaLysGlyAspArgGlyGluThrGlyProAlaGlyProProGlyAlaProGlyAlaPro 1040  
QY 219 AGTAGACCACTCTTCGAGCAAGCTGGTCCCGAGCTTGGGTAGGTGACGATCCACCT 160  
Db 1041 GlyAlaProGlyProValGlyProAlaGlyLys-----AsnGlyAspArg 1055



QY 159 CGTGGGCGCCACCGGGAAGTTGGCGATCTCTCCATCTTCCCGCGCAGAAAGGCGGCA 100  
Db 1056 GlyGluThrGlyProAlaGlyProAlaGlyProIleGlyProAlaGlyAlaArgGlyPro 1075  
QY 99 GCCGCGCCGATCGAAGTCTGCTCTCGAAGT-----61  
Db 1076 AlaGlyProGlnGlyProArgGlyAspLysGlyGluThrGlyGluGlnGlyAspArgGly 1095  
QY 60 -----CCCCCGGGGTGCTGGGGTCT---40  
Db 1096 IleLysGlyHisArgGlyPheSerGlyLeuGlnGlyProProGlySerProGlySerPro 1115  
QY 39 -----CGGCTCGCTCTCGGCA 22  
Db 1116 GlyGluGlnGlyProSerGlyAlaSerGlyProAlaGlyProArgGlyProPro 1133  
RESULT 69  
S23809  
collagen alpha 2(I) chain homolog - sea urchin (Strongylocentrotus purpuratus)  
C:Species: Strongylocentrotus purpuratus (purple urchin)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: S23809  
R:Exposito, J.Y.; d'Alessio, M.; Solursh, M.; Ramirez, F.  
J. Biol. Chem. 267, 15559-15562, 1992  
A:Title: Sea urchin collagen evolutionarily homologous to vertebrate pro-alpha-2(I) coll  
A:Reference number: S23809; MUID:92348411; PMID:1639795  
A:Accession: S23809  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1414 <EXP>  
A:Cross-references: UNIPROT:Q26634; UNIPARC:UPI000007BD61; EMBL:M92040; NID:g161435; PID  
C:Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology  
C:Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix  
F:1207-1414/Domain: fibrillar collagen carboxyl-terminal homology <FCC>  
Alignment Scores:  
Pred. No.: 0.00259 Length: 1414  
Score: 153.00 Matches: 77  
Percent Similarity: 32.9% Conservative: 19  
Best Local Similarity: 26.4% Mismatches: 102  
Query Match: 3.5% Indels: 94  
DB: 1 Gaps: 17  
US-10-768-158-1 (1-2419) x S23809 (1-1414)  
QY 762 CCGGCGCCACGGCGGCTTCAGCGTTCAGCACTGGTCCACAGCTGGTGGCAGTGCT 703  
Db 66 ProGlyProValGlyProGlyProGlyProSerGlyProSerGlyProAlaGlyAsnAsnGly 85  
QY 702 CCGTCAGGCTTCCAGCTGGGCTTGTTCACAGACACACCCCGAGGATC-----655  
Db 86 ProProGly--ProAsnGlyPro-----ArgGlyAsnProGlyMetAspGlyLeuThr 102  
QY 654 -----TGGCGAGTGTCTCCACCTCG-----634  
Db 103 GlyLeuProGlyIleProGlyProProGlyProProGlyLysSerGlySerLeuValAla 122  
QY 633 -----TCACCAGTCCCGATGCATGCTTCATACATCTTGAGAAAAA 595  
Db 123 SerAlaGlnThrSerSerPheAsnLysGlyProSerLeuAlaGlyTyrGlnTyrProGln 142  
QY 594 GCACGTTCGAGTCCATGCGGTGCTCCAGAACTCTCTGCACGTGCTCAACACGAGGACCGT 535  
Db 143 AlaGlnAlaGlyThrProGlyProAlaGlyProPro--GlyProProGlySerArg 161  
QY 534 ACCCGAGCTTATCATTAACCTCCGCGCAGAAATCTTGAAGGTGCTCGGTAGCTCA 475  
Db 162 GlyProGlnGlyLeuThrGlyProSerGlyPro-----172  
QY 474 TGGTCGCGACAGCGGTGGAAGTATGATATAAGACACACACAGATCCTTGGGTTCCGAG 415  
Db 173 ----SerGlyGluThrGlyProSerGlyAsnSerGlyProProGlyProSerGlyLeu---190

QY 414 COATATAGATGACCTTGGAGTCTCCATTTGGAGGTTCAGAGGCGCAGAAAGCGGTAGGGCA 355  
Db 191 ProGlyArg-----ProGlySerAspAspGlyThrPro 203  
QY 354 GGTGGCTCTTGATGAGCGGGGAGAGGTCTCCTTGATGATGTCACGCGCGGTGG 295  
Db 204 GlySerGln-----GlyGlnArg-----GlyProAla---212  
QY 294 GGTACTTCCAGGA-----CCGGAGCTGCTCGTCCA-----265  
Db 213 GlyThrProGlySerArgGlyThrProGlyMetProGlyAlaProGlyMetLysGlyHis 232  
QY 264 -----TGTTTCATCAAGCCGATCTCATCGGGGTGAG-----CGCCCTGGC 226  
Db 233 GlnGlyLeuProGlyMetThrGlySerLysGlyGluArgGlyGluGlyGlyGluArgGly 252  
QY 225 TCACCAAGTAGACCACTCTCGCAGCAAGCTGGTCCGCGACTTGGGTAGGTGACGATCC 166  
Db 253 SerAspGlySerProGlyProValGlyAla-----262  
QY 165 ACACGTCGCTGGGCGCCACCGGGAAGTTGGCGATCTCTCCATCTTCCCGCGCAGAAAG 106  
Db 263 -----ProGlyProAlaGlyProSerGlyGlnProGlyGluArgGlyArgThr 278  
QY 105 GCGGCGCGCGCACCGCATGGAACGCA-----AGTACTTGCTCTCGAACTCCC 58  
Db 279 GlyProAlaGlySerGlnGlyAspArgGlyAlaAspGlyAlaThrGlySerGlnGlyPro 298  
QY 57 CCGGGGTGCTGGGGGTCTCGGCTCGCTCCGCCA 22  
Db 299 ProGlySerThrGly-----ProAlaGlyAlaPro 308  
RESULT 70  
S13141  
hypothetical protein (ribosomal RNA repeat region) - Giardia lamblia  
C:Species: Giardia lamblia  
C:Date: 06-Dec-1996 #sequence\_revision 06-Dec-1996 #text\_change 05-Oct-2004  
C:Accession: S13141; S10886  
R:Upcroft, J.A.; Healey, A.; Mitchell, R.; Boreham, P.F.L.; Upcroft, P.  
Nucleic Acids Res. 18, 7077-7081, 1990  
A:Title: Antigen expression from the ribosomal DNA repeat unit of Giardia intestinalis.  
A:Reference number: S13141; MUID:91088287; PMID:2263466  
A:Accession: S13141  
A:Molecule type: DNA  
A:Residues: 1-294 <UPC>  
A:Cross-references: UNIPROT:O9XZV7; UNIPARC:UPI0000177CC5; EMBL:X52949  
A>Note: the source is designated as Giardia intestinalis  
A>Note: readthrough of the terminator TAG is supposed to occur between residues 241-Ala  
R:Healey, A.; Mitchell, R.; Upcroft, J.A.; Boreham, P.F.L.; Upcroft, P.  
Nucleic Acids Res. 18, 4006, 1990  
A:Title: Complete nucleotide sequence of the ribosomal RNA tandem repeat unit from Giar  
A:Reference number: S10886; MUID:90326542; PMID:2374731  
A:Accession: S10886  
A>Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-241 <HEA>  
A:Cross-references: UNIPARC:UPI0000177CC6; EMBL:X52949  
A>Note: the source is designated as Giardia intestinalis  
A>Note: the assignment of the coding region has been revised in reference S13141  
C:Superfamily: Proline-rich peptide P-B  
Alignment Scores:  
Pred. No.: 0.00219 Length: 294  
Score: 152.00 Matches: 83  
Percent Similarity: 28.8% Conservative: 10  
Best Local Similarity: 25.7% Mismatches: 98  
Query Match: 3.5% Indels: 132  
DB: 2 Gaps: 18  
US-10-768-158-1 (1-2419) x S13141 (1-294)  
QY 759 GGCCCGCGGCGGCGCTTCAGCGTTGCGACTGCTCCACCGCTGGT-----712

Db 12 GlyArgProProGlyArgAlaArgSerGluGluGlyProProThrGlyProGlyProThr 31  
QY 711 GGCAGTGTCTCCGTCAGGGCTTCCAGCTGGCCCTTGTCCAGG----- 670  
Db 32 GlyAlaAlaProGlyAlaValArgAlaProGlyCysSerGlyAlaThrGlyAlaGlnGlu 51  
QY 669 -----ACACCCCCAGGA-----ATCTGGCCAGCT--- 646  
Db 52 AlaAlaGlyGluThrGlyValPheThrGlyProGlyAspArgLeuLeuTrpAlaValArg 71  
QY 645 -----GCTCCACCATCGTCACGAGTCCCGATCGATGCTCTTCAT 607  
Db 72 GlySerPheArgAlaLeuProAlaValProArgSerProAlaPro----- 86  
QY 606 ACTTCAGAAAAAGCAGCTTCGAGTCCATCGCGTGTCTCCAGAACCTCCGACGTGCTCAA 547  
Db 87 -----GlyLeuProSerAlaArgProArgThrTrpAla----- 97  
QY 546 ACCAGGAGCCGTAGCCAGCTTATCATTTATTAACCTCCGGCAGAACTTTGAAAGGTGC 487  
Db 98 -----Cys 98  
QY 486 CTCGGT---AGCTCATGGTCCGACAGAGCGGTGGAACCTGATAATAAGACACCA 433  
Db 99 LeuProValValAlaTrpProAlaGlyGlyAlaGly-----ProProCys 113  
QY 432 -----GATCCTTGGGGTTGCGAGCCATATAGATGACCTTGGAGTCTCCAT 388  
Db 114 ArgProLeuCysLeuAspProGlyArgLeuPro---ArgAlaPro----- 128  
QY 387 TGTGGA-----GCTCAGGGCGAGAAAGCGGTAGGCGAGGT 352  
Db 129 CysGlyProAlaProProProCysAlaHisGlyArgLeuSerValGlnGlyArgProGly 148  
QY 351 GCCTCTTGATGAGCGGGGAGAGTTCAGTTCCT----- 319  
Db 149 GlyGly-----ArgAlaLeuProProGlyHisProLeuGlyHisProAla 163  
QY 318 -----TGATGATGTCAGGCGCGGCT----- 298  
Db 164 LeuProGlnSerSerLeuSerAlaLeuCysLeuSerValCysProAlaProProGlnSer 183  
QY 297 -----GTGGGTACTCCAGACGGGAGCTGCT-----CGTCGATGTTCA 256  
Db 184 LysTrpAlaAlaArgProAlaProArgSerAlaGlyAlaSerLeuArgArgSerGlnGly 203  
QY 255 AGCCGATCTCATCGGGGTACGCGCCCTGCTCACCAGTAGACCACTCTCTGCAGCAAGC 196  
Db 204 SerAlaProTyArgSerAlaAlaProGlyGlyLeuProGlyArgGlyProLeuGlyAsp 223  
QY 195 TGGTCCCGACTTGGGTAGGTGACGATCCACACGTCGCTGGGCGCACCGGAAAGTTGG 136  
Db 224 GlyAlaArgGlyTrpGlyAlaArgProGlyGlyAlaArgGlyGlyValArgAla\*\*\*Arg 243  
QY 135 CQATCTCTCCATCTTCCCGCGCAGAGGGCGGACCGCGCAGCCATCGAAGT 76  
Db 244 ArgSerAlaGlnGlyProArgAlaArgArgGlyAlaAlaArgProGlyAsnArg--- 262  
QY 75 ACTTGCTCTCGAACTCCCGGGGTGCTGGGGTCTCGGCTCTCGCTCT----- 28  
Db 263 -----GlyAlaGlyProArgArgArgSerSerArgThrGly 275  
QY 27 CGCCCATGC 19  
Db 276 ProProCys 278

RESULT 71  
S59856  
collagen alpha 1(iii) chain precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Apr-1996 #sequence revision 19-Apr-1996 #text\_change 09-Jul-2004  
C:Accession: S59856; S62120; S16373

R:Toman, P.D.; de Crombrughe, B.  
Gene 147, 161-168, 1994  
A:Title: The mouse type-III procollagen-encoding gene: genomic cloning and complete DNA  
A:Reference number: S59856; MUID:95011609; PMID:7926795  
A:Accession: S59856  
A:Molecule type: DNA  
A:Residues: 1-1464 <TOM>  
A:Cross-references: UNIPROT:P08121; UNIPARC:UPI0000177386; EMBL:X52046  
R:Toman, D.  
submitted to the EMBL Data Library, November 1994  
A:Reference number: S62120  
A:Accession: S62120  
A:Molecule type: DNA  
A:Residues: 1-866, G', 868-1464 <TOA>  
R:Cross-references: UNIPARC:UPI00000295D6; EMBL:X52046; NID:g575321; PIDN:CMA36279.1; P1  
R:Mesasaeranta, M.; Toman, D.; de Crombrughe, B.; Vuorio, E.  
Biochim. Biophys. Acta 1089, 241-243, 1991  
A:Title: Specific hybridization probes for mouse type I, II, III and IX collagen mRNAs.  
A:Reference number: S16176; MUID:91274355; PMID:2054384  
A:Accession: S16373  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1442-1464 <MET>  
A:Cross-references: UNIPARC:UPI000016CCAB; EMBL:X57983; NID:g50476; PIDN:CMA41048.1; PID  
C:Genetics:  
A:Introns: 29/3; 95/3; 112/3; 150/3; 175/3; 193/3; 211/3; 229/3; 247/3; 265/3; 283/3; 29  
58/3; 673/3; 706/3; 742/3; 760/3; 778/3; 796/3; 814/3; 850/3; 868/3; 886/3; 940/3; 976/3  
C:Superfamily: collagen alpha 1(i) chain; fibrillar collagen carboxyl-terminal homology;  
C:Keywords: coiled coil; extracellular matrix  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-154/Domain: propeptide #status predicted <PRO>  
F:32-92/Domain: von Willebrand factor type C repeat homology <VMC>  
F:155-1464/Product: collagen alpha 1(iii) chain #status predicted <MAT>  
F:1236-1464/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Alignment Scores:  
Pred. No.: 0, 00339 Length: 1464  
Score: 151.50 Matches: 84  
Percent Similarity: 29.5% Conservative: 10  
Best Local Similarity: 26.3% Mismatches: 114  
Query Match: 3.4% Indels: 111  
DB: 2 Gaps: 15

US-10-768-158-1 (1-2419) x S59856 (1-1464)

QY 762 CCCGCCCCACGGGAGGCGCTCAGCTTGACGACCTGTCACCATGTCACCGAGCTGGTGGCAGTGT 703  
Db 892 ProGlyProSerGlyAlaProGlyLysAspGlyProProGlyProAlaGlyAsnSerGly 911  
QY 702 CGGTGAGGGCTTCCA-----GCTGGG---CCTTGTCCACAGG 670  
Db 912 SerProGlyAsnProGlyIleAlaGlyProLysGlyAspAlaGlyGlnProGlyGluLys 931  
QY 669 ACACCCCGCAGGAATCTGGCCAGCTGCTCCACCATCGTCACCAAGT----- 625  
Db 932 GlyProProGly-----AlaGlnGlyProProGlySerProGlyProLeuGlyIleAla 949  
QY 625 -----GCTGGG----- 625  
Db 950 GlyLeuThrGlyAlaArgGlyLeuAlaGlyProProGlyMetProGlyProArgGlySer 969  
QY 624 -----CCCGATCATGCTTCTCATCTTCA 601  
Db 970 ProGlyProGlnGlyIleLysGlyGlyLysProGlyAlaSerGlyHisGlnGly 989  
QY 600 GAAAAAGACGTTCCGAGTCCATCGCTGCCAGAACTCTCCGACGCTGCTCAAACAGG 541  
Db 990 GluArgGlyProProGlyProGlnGlyLeuProGlyGlnProGlyThrAlaGlyGluPro 1009  
QY 540 AGCCGT-----AGCCAGCTTATCATTCATAAACCTCCGCG 505  
Db 1010 GlyArgAspGlyAsnProGlySerAspGlyGlnProGlyArgAspGlySerProGlyGly 1029

```

QY 504 AGAATCTTCTTAAAGGTGCTCGGTAGCTCATGGTCGCAGAGAGCGGTGGAAGTGAAT 445
   ::::
Db 1030 LysGlyAspArgGlyGluAsnGlySer-----ProGlyAlaProGlyAlaProGlyHis 1047
   ::::
QY 444 AAGACACACACAGATCTTTGGGGTTGCGAGCCATATAGATGACCTTGGAGTCTCCATTGT 385
   ::::
Db 1048 ProGlyProProGlyProValGlyProSerGlyLys----- 1059
QY 384 CGAGGTCAGAGCGCAGAAAGCGGTAGGCGAGGTGGCTCTTGTAGAGCGGGAGAGGTCA 325
   ::::
Db 1060 SerGlyAspArgGlyGluThrGlyProAlaGly----- 1070
QY 324 GTTCTCTGATGATGTCACGGCCG-----GCTGTGGGTACTCCAGGACCGGAGCTGCT 271
   ::::
Db 1071 ---ProSerGlyAlaProGlyProAlaGlyAlaArgGlyAlaProGlyProGlnGlyPro 1089
   ::::
QY 270 CQTCGATGTTTCATCAAGCCGATCTCATCGGGTTCAGCGC----- 232
   ::::
Db 1090 ArgGlyAspLysGlyGluThrGlyGluArgGlySerAsnGlyLysGlyHisArgGly 1109
QY 231 ---CCTGGCTCACCAAGTAGACCA---CCTCTCGCAGCAAGCTGGTCCGCACTTTGGGT 178
   ::::
Db 1110 PheProGlyAsnProGlyProProGlySerProGlyAla-AlaGlyHisGlnGlyAla 1129
   ::::
QY 177 AGGTACCATCCACAGTCGCTGGCGCCGACCGGGAAGTTGGCGATCTCTCCATCTTCC 118
   ::::
Db 1129 eGly---SerProGlyProAlaGlyProArgGlyProValGlyPro-----HisGlyPr 1146
   ::::
QY 117 CGCGCAGAGGGCGGAGCGCCACGCCATGGAATCGAAGTACTGCTCTCGAAGTCCC 58
   ::::
Db 1146 oProGly-LysAspGlyThr-----SerGlyHisP 1156
QY 57 CGCGGTCTGGGGTCTCGGCTCGCTCTCGCCATCGCGCGCGTCCGCG 5
   ::::
Db 1156 roGlyProilegly-----ProProGlyPro 1164

RESULT 72
203345
hypothetical protein KIAA0324 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
C:Accession: T02345
R:Ricke, D.O.; Bruce, D.; Mundt, M.; Doggett, N.; Munk, C.; Saunders, E.; Robinson, D.;
re, J.; White, S.; Ueng, S.; Tatam, O.; Campbell, C.; Fawcett, J.; Deaven, L.
submitted to the EMBL Data Library, March 1998
A:Description: Sequencing of human chromosome 16p13.3.
A:Reference number: Z14664
A:Accession: T02345
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1791 <RIC>
A:Cross-references: UNIPROT:O60382; UNIPARC:UPI00000730AE; EMBL:AC004493; NID:g2996648;
C:Genetics:
A:Map position: 16
A:Introns: 1610/2; 1706/2
A>Note: KIAA0324

Alignment Scores:
Pred. No.: 0.00355 Length: 1791
Score: 151.50 Matches: 98
Percent Similarity: 34.6% Conservative: 22
Best Local Similarity: 28.2% Mismatches: 121
Query Match: 3.4% Indels: 106
DB: 2 Gaps: 14

US-10-768-158-1 (1-2419) x T02345 (1-1791)

QY 794 GGTGAAGATGCTTTCCACAGCCCAACTCTTCCCGCCGCGCCAGCGGCTTCAGCGTT 735
   ::::
Db 649 GlySerAspSerSerProGluProLysAla--ProAlaProArgAlaLeuProArgArGs 668
   ::::
QY 734 GCAGCACTGGTCCACCACTGGTGGCACTGCTCGTCAAGGCTTCCACTGGGCGCTTCTC 675
   ::::

```

## RESULT 73

A54849  
 collagen alpha 1(VII) chain precursor - human  
 N:Alternate names: procollagen alpha 1(VII) chain  
 C:Species: Homo sapiens (man)  
 C>Date: 04-Nov-1994 #sequence\_revision 04-Nov-1994 #text\_change 09-Jul-2004  
 C:Accession: A54849; P08044; S16316; I56328; A30296; I84686

```

Db 668 erArgSerGlySerSerSerLysGlyArgGlyProSer-----ProGluGlySerSers 686
QY 674 ACAGACACACCCCGAGAAATCTGGCCAGCTGCTCCACCA----- 637
   ::::
Db 686 erThrGluSerSer-----ProGluHisProProLysSerArgThrAlaArgG 703
   ::::
QY 636 ---TCGTACACAGGTCCCGATGATGCTTTCATCTACTTGAAGAAAGCAGCTTCA 585
   ::::
Db 703 LysArgSerSerProGluProLysThrLysSerArgThrProProArgArgArGsSers 723
   ::::
QY 584 GTCCATCGGTGCTCCAGAACTCTCTGCACGTCT----- 550
   ::::
Db 723 erArg---SerSerProGluLeuThrArgLysAlaArgLysSerArgSerArgSera 742
   ::::
QY 549 ---CAACACAGGAGCCGTAGCCAGCTTATCAT-----TCAT 516
   ::::
Db 742 LaSerSerSerProGluThrArgSerArgThrProProArgHisArgArgSerProServ 762
   ::::
QY 515 AAACCTCGGCGAGAAATCTTGAAGGTGCTCGGTAGTCTATGTCCTCCGAGAGCGGTG 456
   ::::
Db 762 alSerSerProGluProAlaGluLysSerArgSerSerArgArgArgSerAlaSerS 782
   ::::
QY 455 GAACTGATAATAAGACACACCATCTCTGGGGTTGCGAGCCATATAGATGACCTTGA 396
   ::::
Db 782 erProArgThrLys---ThrThrSerArgGly-----Ar 793
QY 395 GTCTCCA-----TTGTGGAGGTACAGGGCGAGAAAGCGTAGGCGAG 354
   ::::
Db 793 gSerProSerProLysProArgGlyLeuGlnArgSerArgSerArgSerArgGluL 813
   ::::
QY 353 GTGGCTCTTGATAGCGGGAGAGGTGAGTCTCTCTTGATGATGCTC----- 308
   ::::
Db 813 sThrArgThrThrArgArgArgAspArgSerGlySerSerGlnSerThrSerArgAr 833
   ::::
QY 307 ---AGGCCGCGCTGTGGGTAC----- 290
   ::::
Db 833 gLlnArgSerArgSerArgSerArgValThrArgArgArgGlyGlySerGlyTyThi 853
QY 289 ---TCCAGGACCGGAGAGCTGCTGCTGATGTTTCAT 258
   ::::
Db 853 sSerArgSerProAlaArgGlnGluSerSerArgThrSerSer--ArgArgArgGlyA 873
QY 257 CAAGCCGATCTATCGGGGTGAGCGCTGCTGCTACCAAGTAGACACCTCTCTGAGCAA 198
   ::::
Db 873 rgSerArgThrProProThrSerArgLysArgSerArgSerArgThrSerProAla--- 891
   ::::
QY 197 GCTGTGCGCGACTTGGGTAGGTGACGATCCACAGTCCGCTGGCGCCGCGGAAGTT 138
   ::::
Db 892 ---ProThrLysArgSerArgSerArg-----AlaSerProAlaThrH 905
   ::::
QY 137 GCGCATCTCTCCATCT-----TCCCGCGCGCAGAAAGG 105
   ::::
Db 905 isArgArgSerArgSerArgThrProLeuLleSerArgArgSerArgSerArgThrS 925
QY 104 CGGAGCGCGCAGCCATGGAACCTGCAAGTACTTGTCTCTCGAACT----- 61
   ::::
Db 925 erProValSerArgArgSerArgSerArgThrSerValThrArgArgSerArgS 945
   ::::
QY 60 ---CCCCCGGGGTGCTGGGGGTCTCGGCTCTCGCTCTCCGCGCATCCCGCCGCC 12
   ::::
Db 945 erArgAlaSerProValSerArgArgSerArgThrProProValThrArgA 965
   ::::
QY 11 GTCGCGCGTCCG 1
   ::::
Db 965 rGArgSerArg 968

```

R;Christiano, A.M.; Greenspan, D.S.; Lee, S.; Uitto, J.  
 J. Biol. Chem. 269, 20256-20262, 1994  
 A;Title: Cloning of human type VII collagen. Complete primary sequence of the alpha1(VII)  
 A;Reference number: A54849; MUID:94327588; PMID:805117  
 A;Accession: A54849  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 1-2944 <CHR>  
 A;Cross-references: UNIPROT:Q02388; UNIPARC:UPI000017A138; GB:L02870; NID:G987124; PIDN:  
 R;Tanaka, T.; Takahashi, K.; Furukawa, F.; Imamura, S.  
 Biochem. Biophys. Res. Commun. 183, 956-963, 1992  
 A;Title: Molecular cloning and characterization of type VII collagen cDNA.  
 A;Reference number: PH0844; MUID:922331902; PMID:1567409  
 A;Accession: PH0844  
 A;Molecule type: mRNA  
 A;Residues: 'EPR', 340-475, 'RALSTASHSTLCWRATRWHPNCRGSHWTRAAACEPCNRFPASHRAARAG', 524-528, 'C',  
 A;Cross-references: UNIPARC:UPI000017A139; DDBJ:D11152; DDBJ:D13694; NID:G453698; PIDN:E  
 A;Experimental source: keratinocyte  
 A;Note: the authors translated the codon ACC for residues 394 and 397 as Tyr  
 R;Parente, M.G.; Chung, L.C.; Ryyanen, J.; Woodley, D.T.; Wynn, K.C.; Bauer, E.A.; Mac  
 Proc. Natl. Acad. Sci. U.S.A. 88, 6931-6935, 1991  
 A;Title: Human type VII collagen: cDNA cloning and chromosomal mapping of the gene.  
 A;Reference number: S16316; MUID:91334380; PMID:1871109  
 A;Accession: S16316  
 A;Molecule type: mRNA  
 A;Residues: 815-892, 'E', 894-1439 <PAR>  
 A;Cross-references: UNIPARC:UPI000016A722; GB:M65158; GB:S49017; NID:G180914; PIDN:AAA96  
 A;Experimental source: keratinocyte  
 R;Gannon, W.R.; Abernethy, M.L.; Padilla, K.M.; Priasayanh, P.S.; Cook, M.E.; Wright, J.;  
 J. Invest. Dermatol. 99, 691-696, 1992  
 A;Title: Noncollagenous (NC1) domain of collagen VII resembles multidomain adhesion prot  
 A;Reference number: I56328; MUID:93107742; PMID:1469284  
 A;Accession: I56328  
 A;Status: translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 'EPR', 372-517, 'DV', 520-540, 'W', 542-1255 <RES>  
 A;Cross-references: UNIPARC:UPI000016B3AC; GB:S51236; NID:G262308; PIDN:AAB24637.1; PID:  
 R;Setzler, J.L.; Eisen, A.Z.; Bauer, E.A.; Morris, N.P.; Glanville, R.W.; Burgeson, R.E.  
 J. Biol. Chem. 264, 3822-3826, 1989  
 A;Title: Cleavage of type VII collagen by interstitial collagenase and type IV collagena  
 A;Reference number: A30296; MUID:89139437; PMID:2537292  
 A;Accession: A30296  
 A;Molecule type: protein  
 A;Residues: 'A', 1240-1246, 'G', 1248-1250, 'XE', 1253-1255, 'Q', 1257, 'E', 2032, 'C', 2034-2041;  
 A;Cross-references: UNIPARC:UPI000017A13A; UNIPARC:UPI000017A13B; UNIPARC:UPI000017A13C;  
 A;Note: two reported peptides cannot be reliably located  
 R;Greenspan, D.S.  
 Hum. Mol. Genet. 2, 273-278, 1993  
 A;Title: The carboxyl-terminal half of type VII collagen, including the non-collagenous  
 A;Reference number: I48103; MUID:93271985; PMID:8499916  
 A;Accession: I48103  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 2395-2871, 'S', 2873-2944 <RES>  
 A;Cross-references: UNIPARC:UPI000016A724; GB:L06862; NID:G388713; PIDN:AAA89196.1; PID:  
 R;Christiano, A.M.; Ryyanen, M.; Uitto, J.  
 Proc. Natl. Acad. Sci. U.S.A. 91, 3549-3553, 1994  
 A;Title: Dominant dystrophic epidermolysis bullosa: identification of a Gly --> Ser subs  
 A;Reference number: A55255; MUID:94224777; PMID:8170945  
 A;Content: annotation  
 C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
 ed and subsequently O-glycosylated.  
 C;Genetics:  
 A;Gene: GDB:COL7A1; EBR1; EBD1; EB  
 A;Cross-references: GDB:128750; OMIM:120120  
 A;Map position: 3p21.3-3p21.3  
 A;Note: defects in this gene can result in dominant and recessive dystrophic epidermolys  
 A;Note: there are 118 introns  
 C;Complex: type VII collagen is probably a homotrimer  
 C;Function:  
 A;Description: structural component of extracellular polymer associated with anchoring f  
 C;Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproli  
 F.1-16/Domain: signal sequence #status predicted <SIG>

F.17-2944/Product: collagen alpha 1(VII) chain #status predicted <MAT>  
 F.17-1253/Domain: amino-terminal nonhelical #status predicted <NC1>  
 F.36-201/Domain: von Willebrand factor type A repeat homology <VWA1>  
 F.231-318/Domain: fibronectin type III repeat homology <FN1>  
 F.327-413/Domain: fibronectin type III repeat homology <FN2>  
 F.414-502/Domain: fibronectin type III repeat homology <FN3>  
 F.508-593/Domain: fibronectin type III repeat homology <FN4>  
 F.598-683/Domain: fibronectin type III repeat homology <FN5>  
 F.686-771/Domain: fibronectin type III repeat homology <FN6>  
 F.776-862/Domain: fibronectin type III repeat homology <FN7>  
 F.864-952/Domain: fibronectin type III repeat homology <FN8>  
 F.954-1045/Domain: fibronectin type III repeat homology <FN9>  
 F.1052-1219/Domain: von Willebrand factor type A repeat homology <VWA2>  
 F.1170-1172/Region: cell attachment (R-G-D) motif  
 F.1189-1253/Region: cysteine/proline-rich  
 F.1254-2783/Region: interrupted helical  
 F.1334-1336/Region: cell attachment (R-G-D) motif  
 F.2008-2010/Region: cell attachment (R-G-D) motif  
 F.2553-2555/Region: cell attachment (R-G-D) motif  
 F.2784-2944/Domain: carboxyl-terminal nonhelical #status predicted <NC2>  
 F.2876-2929/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>  
 F.337,786,1109/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F.2167,2176,2185,2188,2664,2667,2673/Modified site: 4-hydroxyproline (Pro) #status exper  
 F.2625,2631/Modified site: 5-hydroxylysine (Lys) #status experimental  
 F.2625,2631/Binding site: carbohydrate (Lys) (covalent) #status experimental  
 F.2634,2802,2804/Disulfide bonds: interchain #status predicted

Alignment Scores:  
 Pred. No.: 0.00431 Length: 2944  
 Score: 151.00 Matches: 83  
 Percent Similarity: 33.8% Conservative: 15  
 Best Local Similarity: 28.6% Mismatches: 126  
 Query Match: 3.4% Indels: 66  
 DB: 2 Gaps: 15

US-10-768-158-1 (1-2419) x A54849 (1-2944)

Qy	762	CCCGCCCGCCAGGGCCCTCAGCGTTCAGCAGCTGGTCCACGAGCTGGTGGCAGTGCT	703
Db	2185	ProGlyProProGlyAlaPro-----GlyLeuAlaGlyProAlaGlyProGlnGly	2201
Qy	702	CCGTCAGGGCTTCCAGCTGGGCTTGTCCACAGGACACCCCGAGGAATCTGCCAGCTGCT	643
Db	2202	ProSerGlyLeuLysGlyGluProGlyGluThrGlyProGlyArgGlyLeuThrGly	2221
Qy	642	CCA-----CCATCGTCACAGTCCCGATGCGATGCTTCATCACTGA	601
Db	2222	ProThrGlyAlaValGlyLeuProGlyProGlyProSerGlyLeuVal-----	2238
Qy	600	GAAGAACACGTTCCAGTCCATGCGTCTCCAGAACTCTCTGCACGT---GCTCAAGC	544
Db	2239	-----GlyProGlnGlySerProGlyLeuProGlyGlnValGlyGluThr	2253
Qy	543	AGGAGCCGTAGCCAGCTTATCATTAACCTCCGCGCAGAAATCTTTGAAAGTGCCTC	484
Db	2254	GlyLysProGlyAlaProGlyArgAspGlyAlaSerGlyLysAspGlyArg-----	2271
Qy	483	GGTACTCATGTCGCGAGAGCGGTGGAAGTGAATAGACACCACTCTTGG	424
Db	2272	GlySerProGlyValProGlySerProGlyLeu-----ProGlyProVal	2286
Qy	423	GTTTCGAGCCATATAGATGACCTGGAGTCTCCATTTGGAGGTTCAGAGGCGAAGAC	364
Db	2287	Gly-----ProLysGlyGluProGlyPheThrGlyAlaProGlyGln---AlaVal	2303
Qy	363	GGTAGGCGAGGTGGCTTCTTGATGAGCGGGGAGAGTCACTTCTTGATGTCCAGGC	304
Db	2304	GlyLeuProGlyAlaLys-----GlyGluLysGlyAlaProGlyGlyLeuAlaGly	2320
Qy	303	CCGCTGTGGGTACTCCAGGACCGGAGCTGCTCGATGTTTCATCAAGCGCATCTCAT	244
Db	2321	AspLeuValGlyGluProGlyAlaLysGlyAspArgGlyLeuProGlyProArgGlyGlu	2340

QY 243 CGGGTCAG-----CGCCCTGGCTCACCAGTAGACACCTCCCTGCACCAAGCTGGTGC 190  
 Db :|||:|||||  
 2341 LysGlyGluAlaGlyArgAlaGlyGluProGlyAspProGlyGluGly-GlnLysGlyAl 2360  
 QY 189 C-----GAGCTTGGGTAGGTGACGATCCACA 163  
 Db :|||:|||||  
 2360 aProGlyProLysGlyPheLysGlyAspAspProGlyValGlyValProGlySerProGl 2380  
 QY 162 CTTCTGGGCGCCACCGGGAAGTTGGCATCTCTCCATC----- 122  
 Db :|||:|||||  
 2380 yProProGlyProProGlyValLysGlyAspLeu-GlyLeuProGlyLeuProGlyAlaP 2400  
 QY 121 -----TTCCCGCGCAGAGGCG-----GGCAGCCGACGC 91  
 Db :|||:|||||  
 2400 roGlyValValGlyPheProGlyGlnThrGlyProArgGlyGluMetGlyGlnProGlyP 2420  
 QY 90 CATGGAACCTCGAAGTACTTGCTCTCGAACTCCCGGGGTGCTGGGGTCTCGGCTCGC 31  
 Db :|||:|||||  
 2420 ro--SerGlyGluArgGlyLeuAlaGlyProProGlyArgGluGlylleProGlyProL 2439  
 QY 30 TCTCCGCCATGCCCGCGCGCTCGCGC 5  
 Db :|||:|||||  
 2439 euGly-----ProProGlyPro 2444

RESULT 74  
 149607  
 procollagen type V alpha 2 - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
 C:Accession: I49607  
 R.Andrikopoulos, K.; Suzuki, H.R.; Solursh, M.; Ramirez, F.  
 Dev. Dyn. 195, 113-120, 1992  
 A:Title: Localization of pro-alpha 2(V) collagen transcripts in the tissues of the devel  
 A:Reference number: I49607; PMID:93214071; PMID:1297453  
 A:Accession: I49607  
 A>Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1497 <RES>  
 A:Cross-references: UNIPROT:Q61431; UNIPARC:UPI0000028657; GB:L02918; NID:g309180; PIDN:  
 C:Genetics:  
 A:Gene: Col5a-2  
 C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;  
 F:39-98/Domain: von Willebrand factor type C repeat homology <WMC>  
 F:1270-1497/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Alignment Scores:  
 Pred. No.: 0.00406 Length: 1497  
 Score: 150.50 Matches: 83  
 Percent Similarity: 30.1% Conservative: 11  
 Best Local Similarity: 26.6% Mismatches: 123  
 Query Match: 3.4% Indels: 95  
 DB: 2 Gaps: 16

US-10-768-158-1 (1-2419) x I49607 (1-1497)

QY 783 CTTTCCACAGCCCACTCTCCCGGCCCA-----CGGGCAGGCGCTCAGCGTTGCAGC 730  
 Db :|||:|||||  
 902 ProProGlyAlaThrGlyPheProGlySerAlaGlyArgValGlyProPro----- 918  
 QY 729 ACTGGTCCACAGCTGGTGGCAGTCTCGTCAGAGGCTCCAGCTGGG---CCTTGTAC 673  
 Db :|||:|||||  
 919 -----GlyProAlaGlyAlaProGlyProAlaGly---ProAlaGlyGluProGlyLys 935  
 QY 672 AGGACACACCCAGGAATC-----TGG 652  
 Db :|||:|||||  
 936 GluGlyProProGlyLeuArgGlyAspProGlySerHisGlyArgValGlyAspArgGly 955  
 QY 651 CCAGCTGTCTCCACATCGTCACCAAGTCCCGATGCTTCATCATTTGAGAAAAGCA 592  
 Db :|||:|||||  
 956 ProAlaGlyProProGlySerProGlyAspLysGly-----AspProGlyGluAspGly 973  
 QY 591 CTTTCAGTCCATGGTCTCTCCAGAACTCTCTGCAGCTGCTCAAAACAGGAGC----- 538  
 Db :|||:|||||

Db 974 GlnProGlyProAspGlyProProGlyProAlaGlyThrGlyGlnArgGlylleVal 993  
 QY 537 -----CGTAGCCCACTTATCAT 520  
 Db 994 GlyMetProGlyGlnArgGlyValThrGlyMetProGlyLeuProGlyProAlaGly--- 1012  
 QY 519 TCATAAACTCCGCGCAGAAATCTTCAAAGTGCCTCGGTAGCTCATGTCGCCACAGAGC 460  
 Db 1013 -----ThrProGlyLysVal-----GlyProThr 1020  
 QY 459 GGTGAACCTGATAAATAAGACACACCAGATCCTTGGGTTGCGAG-----CCA 412  
 Db 1021 GlyAlaThrGlyAspLysGlyProProGlyProValGlyProProGlySerAsnGlyPro 1040  
 QY 411 TATAGACACCTTGGAGTCTCCATTGTGAGAGTACAGAGGCGCAGAAAGCGGTAGGACAGT 352  
 Db 1041 ValGlyGluProGlyProGluGlyProAlaGlyAsnAspGlyThrProGlyArgAspGly 1060  
 QY 351 GGCTCTTGATAGCGCGGAGAGGTTCAGTCTCTTGATGATGTCACAGGCCG----- 301  
 Db 1061 AlaValGlyGluArgGlyAspArgGly-----AspProGlyProAlaGlyLeu 1076  
 QY 300 -----GCTGTGGTACTCCAGGACCGGAGCTGCTCGTATGTCATCA 256  
 Db 1077 ProGlySerGlnGlyAlaProGlyThrProGlyProValGlyAlaProGlyAspAlaGly 1096  
 QY 255 AGCCGATCTCATCGGGGTGAGCGCTCGTCCACCAAGTAGACACCTCTCTCCAGCAAGC 196  
 Db 1097 GlnArgGlyGluProGlySerArg-----GlyProValGlyProProGlyArgAla--- 1113  
 QY 195 TGGTCCCGACTTGGGTTAGTGACGATCCACACGTCGCTGGCGCCGACCGGAAGTTGG 136  
 Db 1114 -----GlyLysArgGlyLeuProGlyProGlnGlyPro 1124  
 QY 135 CGATCTCTCCATCTTCCCGCGGAGAAAGCGCGCAGCGCATCGAACTCGAACT 76  
 Db 1125 ArgGlyAspLysGlyAspAsnGlyAspArgGlyAspArgGlyGlnLysGlyHisArgGly 1144  
 QY 75 ACTTGCTCTCGAACT-----CCCGCGGGTCTGGGGTCTCGG----- 37  
 Db 1145 PheThrGlyLeuGlnGlyLeuProGlyProGlyProAsnGlyGluGlnGlySerAla 1164  
 QY 36 -----CCTCGCTCTCGGCCA 22  
 Db 1165 GlylleProGlyProPheGlyProArgGlyProPro 1176

RESULT 75  
 T45025  
 mucin MUC5B, tracheobronchial [imported] - human (fragment)  
 C:Species: Homo sapiens (man)  
 C>Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000  
 C:Accession: T45025  
 R.Desseyn, J.L.; Guyonnet-Duperat, V.; Porchet, N.; Aubert, J.P.; Laine, A.  
 J. Biol. Chem. 272, 3168-3178, 1997  
 A:Title: Human mucin gene MUC5B, the 10.7 kb large central exon encodes various alternat  
 A:Reference number: 222899; PMID:97166151; PMID:9013550  
 A:Accession: T45025  
 A>Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-3570 <DES>  
 A:Cross-references: UNIPARC:UPI0000016AD88; EMBL:272496; NID:g1834502; PIDN:CAA96577.1; P  
 C:Experimental source: placenta  
 C:Genetics:  
 A:Gene: MUC5B

Alignment Scores:  
 Pred. No.: 0.00491 Length: 3570  
 Score: 150.50 Matches: 206  
 Percent Similarity: 29.8% Conservative: 83  
 Best Local Similarity: 21.2% Mismatches: 326  
 Query Match: 3.4% Indels: 357  
 DB: 2 Gaps: 42



QY	600	GRAAAAGCAGCTTCGAGTCCATCGGTGCTCCAGAACTCCTCCAGACGTCTCAAAACGAGG	541
Db	2180	GlyThrSerArgThrThr-----AlaThrAlaThrProSerLysThrArgThrSer	2196
QY	540	AGCGGTAGCCAGCTTATCATTAACCTCCGCGCAGAAATCTT-----	496
Db	2197	ThrLeuLeuProSer-----SerProThrSerAlaProIleThrThrValValThrThr	2214
QY	495	-----GAAAGTGCTCGGTGATCATCTGATGCTCCGACAGAGCGGTGAACTGATAAT	445
Db	2215	GlyCysGluProGlnCysAlaTrpSerGluTrp-----LeuAspTyr	2228
QY	444	AGACACCA-----CCAGATCTTGGGT-----	421
Db	2229	SerTyrProMetProGlyProSerGlyGlyAspPheAspThrTyrSerAsnIleArgAla	2248
QY	420	-----TGCAGCCATATAGATGACCTTGAGTCTCCATTGTGGAGTCCAG	376
Db	2249	AlaGlyGlyAlaValCysGluGln-----ProLeuGlyLeuGluCysArgAlaGln	2265
QY	375	AGGCGAGAAAGCGGTA-----GGGCAGGTG-----GCTCTTGAT--	342
Db	2266	-AlaGlnProGlyValProLeuArgGluLeuGlyGlnValValGluCysSerLeuAspPh	2285
QY	341	-----CAGCGGGGAGAGGTGAGTTCCTTGTATGATGCCA	307
Db	2285	eglyLeuValCysArgAsnArgGluGlnValGlyLysPheLysMetCysPheAsnTyrGl	2305
QY	306	GGCCCGCGTGGGTACTCCAGACCGGAGCTGC-----TCGTCGATGT	262
Db	2305	uileArgValPheCys-CysAsnTyrGlyHisCysProSerThrProAlaThrSerSerT	2325
QY	261	TCATCAAGCCGATCTCATCGGGGTACGCCCTCG--CTCACCAAGTAGACCACTCCT	205
Db	2325	hAlaThrProSerSerThrProGlyThrThrTrpIleLeuThrLysLeuThrThrThr-	2344
QY	204	GCAGCAAGCTGTGCGGACTGGGGTAGTGACGATCCACAGTCGTCGGGCGCACCG	145
Db	2345	AlaThrThrThr-----GluSerThrGlySerThrAlaThrPro	2357
QY	144	GGAAGTTGGCGATCTCTCCATCTTCCCGCGCAGAAAGGGCGCAGCCGACCCATGGA	85
Db	2358	SerSerThrGlnGlyProProAlaGlyThrProHisValSerThrThrAlaThrThrPro	2377
QY	84	ACTCGAAGTACTGCTCTCGAACTCCC-----CCGGGGTGTGGGGGTCTCGGCCT	34
Db	2378	ThrValThrSerSerLysAlaThrProPheSerSerProGlyThrAlaThrAlaLeuPro	2397
QY	33	CGCTCTCGCCATGCGGCC	15
Db	2398	-AlaLeuArgSerThrAla	2403
RESULT 76			
T20906			
hypothetical protein F14F7.1 - Caenorhabditis elegans			
C:Species: Caenorhabditis elegans			
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004			
C:Accession: T20906			
R:McMurray, A.			
submitted to the EMBL Data Library, November 1996			
A:Reference number: Z19345			
A:Accession: T20906			
A:Status: preliminary; translated from GB/EMBL/DBJ			
A:Molecule type: DNA			
A:Residues: 1-305 <WIL>			
A:Cross-References: UNIPROT:O17805; UNIPARC:UPI0000077AA4; EMBL:Z81503; PIDN:CAB04111.1;			
C:Experimental source: clone F14F7			
C:Genetics:			
A:Gene: CESP:F14F7.1			
A:Map position: 3			
A:Introns: 27/3; 49/3			
Alignment Scores:			

RESULT 77

T32248

hypothetical protein T15B7.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

Pred. No.: 0.00313 Length: 305  
Score: 150.00 Matches: 70  
Percent Similarity: 26.3% Conservative: 5  
Best Local Similarity: 24.6% Mismatches: 78  
Query Match: 3.4% Indels: 132  
DB: 2 Gaps: 15

US-10-768-158-1 (1-2419) x T20906 (1-305)

QY 774 GCCAACTTCCCGCGCCACCGGCGCTCAGCGTTCGACGACT----- 727

Db 87 AlaGlnProGlnGlyGlyGlyGlyGlyGlyGlnCysSerThrCysCysArgPro 106

QY 726 GTCCACCAAGCTGGTGGCAGTCTCGTCA----- 697

Db 107 GlyProProGlyAlaGlyGlyThrProGlyLysProGlyArgProGlyAlaProGlyAla 126

QY 696 ---GGCTTCCA-----GCTGGCGCTTGTAC-----AGG 670

Db 127 TrpGlyMetProGlyAsnProGlyLysGlyGlySerGlyProCysHisProValValPro 146

QY 669 ACACCC-----CCAGGAATCTGCCAGCTGCTCCACCATCG 634

Db 147 ThrProCysLysProCysProGlyGlyArgProGly-----ProProGlyProProGly 164

QY 633 TCACCAAGTCCGATGCATGCTTCTCATACTTGAGAAAAGCACGTTTCGAGTCCATCGGT 574

Db 165 ArgProGlySerAsp-----Gly 170

QY 573 GCTCCAGNACTCTCCAGCTGCTCAACCAGAGCCGTAGCCAGCTTATCATTTCTATAA 514

Db 171 GlnProGlyArgProAla----- 176

QY 513 ACTCCGGCAGAAATCTTGAAGGTGCTCGGTAGTCTCATGTCGCGCAGAGACGCGTGA 454

Db 177 ThrGlyGly-----GlyAla 181

QY 453 ACTGATAATAAGACACCAAGATCCTTGGGGTTCGAGCCATATAGATGACCTTGGAGT 394

Db 182 AlaProArgProGlyProProGlyProLysGlyProArg----- 194

QY 393 CTCCATTGTGGAGTCAGAGGCAGAAAGCGTAGGCGAGGTGGCTCTTGATGAGCGCGG 334

Db 195 -----GlyAlaProGlyAsnSerGlyArgAlaGly----- 204

QY 333 GAGAGTTCAGTTCCTTGATGATGTCAGGCCCG----- 301

Db 205 -----AlaProGlyGlnProGlyAsnAspAlaHisGlyTyrGly 217

QY 300 ---GCTGTGGTACTCCAGACCGGAGCTGCTCGATGTTTCATCAAGCCGATCTCAT 244

Db 218 GlyGlyValGlyAlaProGlyProAlaGlyProArgGlyAlaProGlyProAlaGlyHis 237

QY 243 CGGGGTACGCCCTTGGCTCACCAGTAGACACCACTCTTCGAGCAGAGTGGTCCGGACT 184

Db 238 ProGlySerSerGlyGly-----GlyArgProGlyProAlaGlyPro----- 251

QY 183 TGGGGTAGGTGACGATCCACAGTCGCTGGGCGCCAGCGGGAAGTTGGCGATCTCCT--- 127

Db 252 -----LysGlyAlaProGlyGlnProGlyGlnProGlyProGly 262

QY 126 -----CCATCTTCCGCGCAGAGGCGCGCAGCGCATGCGCAACTCGA 79

Db 263 ProAspGlyHisProGlyGlnProGlyArgProGlyGlnSerGlyGlySerGlyAsnArg 282

QY 78 AGTACTTCTCTCGA 64

Db 283 GlyValCysProLys 287

C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T32248  
R;Pauley, A.; Gattung, S.  
submitted to the EMBL Data Library, September 1997

A/Description: The sequence of C. elegans cosmid T15B7.  
A/Reference number: Z21139

A/Accession: T32248  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-325 <PAU>

A/Cross-references: UNIPROT:O17036; UNIPARC:UPI000007698C; EMBL:AF022985; PIDN:AAB69961

A/Experimental source: strain Bristol N2; clone T15B7

C/Genetics:  
A/Gene: CESP:T15B7.4  
A/Map position: 5  
A/Introns: 266/1

Alignment Scores:					
Pred. NO.:	0.00347	Length:	325		
Score:	149.50	Matches:	68		
Percent Similarity:	32.1%	Conservative:	18		
Best Local Similarity:	25.4%	Mismatches:	79		
Query Match:	3.4%	Indels:	103		
DB:	2	Gaps:	11		

US-10-768-158-1 (1-2419) x T32248 (1-325)

QY	732	AGCACTGGTCCACCAAGCTGGTGGCAGTCTCGTCAGGGCTTCCAGCTGGCCCTGTGTAC	673
DB	86	SerArgGlyCysProAlaGlyProProGlyProGlyGlnProGlyAlaProGlyGlu	105
QY	672	AGGACACCCCCAGGAATC-----	655
DB	106	GlnGlyHisProGlyLeuAlaGlyGlnProGlySerGlyAlaArgIleAsnProAlaThr	125
QY	654	-----	637
DB	126	GlyArgProGlyPheCysIleThrCysProAlaGlyAlaProGlyProAlaGlyProPro	145
QY	636	TCGTCCACCAAGTCCCGATGCATGCTTCACTTGCAGAAAAGCAGCTTCGAGTCCATGC	577
DB	146	GlyAlaProGlyPro-----	157
QY	576	GGTGTCTCCAGAACTCTCTGCACGTCTCAAAACAGGAGCCGTAGCCAGCTTATCAITCA	517
DB	158	GlyAlaPro-----	169
QY	516	TAAACCTCCGCAGAAATCTTTGAAAGGTGCCTCGGTAGCTCATGGTCGCAGAGAGCGGT	457
DB	170	-----	178
QY	456	GGAACGTGATAATAGACACCAACCAGATCCTTTGGGGTTCCGAGCCATATAGATGACCTTGG	397
DB	179	GlySerProGlyGln---ProGlyHisProGlySer-----	191
QY	396	AGTCTCCATTGTGGAGGTACAGAGGCAGAA-----	352
DB	192	AsnProGlyArgGlyGlyGlnArgSerArgGlyLeuProGlyProSerGlyArgProGly	211
QY	351	GGCTCTTGATGAGCGGGGAGAGGTTCAGTTCTCTTGATGATGTCCAGGCCCGCGCT-----	298
DB	212	Pro-----	218
QY	297	-----	265
DB	219	ProGlyGlnProGlyHisSerGlyGlyAlaGlySerProGlyProGlyGlnProGlyProGly	238
QY	264	TGTTTCATCAAGCCGATCTCATCGGGGTACAG-----	220
DB	239	ProSerGlyGlnProGlyHisSerGlyAsnAspGlyValProGlyAlaProGlyAsnPro	258
QY	219	AGTAGACCAACCTCTCTGCAGCAAGCTGGTCCCGGACTTGGGGTAGTGACGATCCACACGT	160



```
Db 152 Val-----GlyProAlaGlyProAlaGlyAspGlnGlyArgHis 164
QY 441 ACACCACCATCTTGGGTGGAG-----CCATATAGATGACCTTGGAGTCTCCAT 388
Db 165 GlyProGlyProThrGlyGlyGlnGlyPro----- 176
QY 387 TGTGGAGGTGAGAGGCGAGAAAGCGGTAGGCAGGTGGCTTTGATGAGCGGGGAGAG 328
Db 177 -----GlyGluGlnGlyAspAlaGlyArgProGlyAlaAla----- 188
QY 327 TCAGTTCTTCATGATGATCCAGGCCCGGCTGTGGGTACTCCAGGACCGGAGCTGCTCGT 268
Db 189 -----GlyCysProGlyPro-----ProGlyPro----- 196
QY 267 CGATGTTTCATCAAGCCGATCTATCGGGGTGTCAGGCCCTGGCTCACAAGTAGACCACT 208
Db 197 -----ArgGlyGlu-----ProGlyThrGluTyArgProGly 207
QY 207 CCTGCAGCAAGCTGGTCCGGACTTGGGGTAGGTGACGATCCACACGTCCCTGGGCGCA 148
Db 208 GlnAla-----GlyArg-----AlaGly 213
QY 147 CCGGGAAGTTGGGATCTCTCCATCTTCCCGCGGCAGAAAGCGGCGCAGCCGACCAT 88
Db 214 ProProGlyProArgGlyProProGlyProGluGlyAsnProGlyGlyAlaGlyGluAsp 233
QY 87 GGAACTCGAAGTACTTGCTCTCGAATCCCGCGGGGTGCTGGGGTCTCGGCTCGCTCT 28
Db 234 GlyAsnGlnGlyProValGlyHis-----ProGlyValProGly---ArgProGlyIle 250
QY 27 CCG 25
Db 251 Pro 251

RESULT 79
T43481
probable mucin DKFPz434C196.1 - human (fragment)
N:Alternate names: protein DKFPz434B0635.1
C:Species: Homo sapiens (man)
C>Date: 21-Jan-2000 #sequence revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T43481; T34549; T17264
R:Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, December 1999
A:Reference number: Z22514
A:Accession: T43481
A:Molecule type: mRNA
A:Residues: 1-580 <AAA>
A:Cross-references: UNIPROT:Q9UF83; UNIPARC:UPI000006D400; EMBL:AL133561; NID:G6599133;
A:Experimental source: adult testis; clone DKFPz434C196
R:Poustka, A.; Weillenreuther, R.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, October 1999
A:Reference number: Z21540
A:Accession: T34549
A:Molecule type: mRNA
A:Residues: 262-580 <POU1>
A:Cross-references: UNIPARC:UPI0000070F36; EMBL:AL117481; NID:G6102864; PIDN: CAB59245.2;
A:Experimental source: adult testis; clone DKFPz434B0635
R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A:Reference number: Z18723
A:Accession: T17264
A:Molecule type: mRNA
A:Residues: 262-580 <POU2>
A:Cross-references: UNIPARC:UPI0000070F36; EMBL:AL117481; NID:G59111958; PIDN: CAB59594.1;
A:Experimental source: adult testis; clone DKFPz434B061
C:Genetics:
A>Note: DKFPz434C196.1; DKFPz434B0635.1; DKFPz434B061.1
```

Alignment Scores:  
Pred. No.: 0.00469 Length: 580  
Score: 148.50 Matches: 76  
Percent Similarity: 33.1% Conservative: 17  
Best Local Similarity: 27.0% Mismatches: 85

```
Query Match: 3.4% Indels: 103  
DB: 2 Gaps: 14  
US-10-768-158-1 (1-2419) x T43481 (1-580)  
QY 789 AGATGTCCTTCCACAGCCCAACTCTTC-----CCCGGCCACGGGCGAGG----- 745  
Db 71 ArgThrProThrArgAlaSerLeuMetArgThrProProArgAlaSerProThrArgLys 90  
QY 744 CCTCAGCGTTGCAGCAGCTGGTCCACAGCTGGTGGCAGT-----GTCCTCGTCA 697  
Db 91 ProProArgAlaSerProArgThrProSerArgAlaSerProThrArgArgLeuProArg 110  
QY 696 GGGCTTCAGCTGGGCTTGTCTAC-----AGACACCCCCAGGAATCTGG 652  
Db 111 AlaSerProMetGlySerProHisArgAlaSerProMetArgThrProProArgAlaSer 130  
QY 651 CCAGCTGCTCCACATCGTCACAGCTCCGAGTCCGATGTCATCTTACATCTTGAGAAAAGCA 592  
Db 131 ProThrGlyThrProSerThrAlaSerPro-----ThrGlyThrProSer 145  
QY 591 CGTTCAGTCCATGCGGTGCTCCC-----AGAACTCTGCACGCTGT 550  
Db 146 SerAlaSerProThrGlyThrProProArgAlaSerProThrGlyThrProProArgAla 165  
QY 549 CAAACAGGAGCGGTAGCCAGCTTATCATTTATAAACCTCCGGCAGAAATCTTGAAGG 490  
Db 166 TrpAla---ThrArgSerProSerThrAlaSerLeuThr-----ArgThrProSerArg 182  
QY 489 TGCCTCGGTGCTCATGGTCCGCGAGAGCGGTGGAACCTGATTAATAAGACACACCCAGAT 430  
Db 183 AlaSerLeuThrArgTrpProProArgAlaSerProThr-----ArgThrProPro--- 199  
QY 429 CTTTGGGGTTGCGAGCCATATAGATGACCTTGGAGTCTCCATTGTGGAGTCCAGAGGCA 370  
Db 199 ----- 199  
QY 369 GAAAGCGGTAGGCGAGGTGGCTTTGATGAGCGGGGAGAGTTCAGTTCCTTGATGATGT 310  
Db 199 ----- 199  
QY 309 CCAGGCCCGGTGTGGGTACTCCAGGACCGGGAGCTGCTCGTGCATGTTTCATCAAGCCGA 250  
Db 200 -----ArgGluSerProArgMet 205  
QY 249 TCTCATCGGGGTGAGCGCCCTGGCTCACCAGATAGACACCTCTCTCGAGCAAGCTGTGC 190  
Db 206 SerHisArg-----AlaSerProThrArgThrProProArgAlaSer----- 219  
QY 189 CGGACTTGGGTAGTGTGACGATCCACAGCTCGTGGGCGCGCACCCGGGAAAGTTGGCGATCT 130  
Db 220 -----ProThrArgArgProProArgAlaSerProThrArgThr 232  
QY 129 CTTCCATCTTCCCGGGGAGAG-----AGGGCGGAGCGCCAGCCCATGGAACCTCGA 79  
Db 233 ProProArgGluSerLeuArgThrSerHisArgAlaSerProThrArgMetProProArg 252  
QY 78 AGTACTTGTCTCTCGAACTCCCGGGGTGC-----TGGGGGTCTCGGCTCGCTCTCCG 25  
Db 253 AlaSerProThrArgArgProProArgAlaSerProThrGlySerProProArgAlaSer 272  
QY 24 CCA 22  
Db 273 Pro 273
```

RESULT 80  
JQ0317  
hypoetical 82K protein - Xanthomonas campestris pv. vesicatoria  
C:Species: Xanthomonas campestris pv. vesicatoria  
C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004  
C:Accession: JQ0317  
R:Bonas, U.; Stall, R.B.; Staakawicz, B.  
Mol. Gen. Genet. 218, 127-136, 1989

A:Title: Genetic and structural characterization of the avirulence gene avrBs3 from Xanth

A:Reference number: JQ0316; MUID:89384426; PMID:2550761

A:Accession: JQ0317

A:Molecule type: DNA

A:Residues: 1-784 <BON>

A:Cross-references: UNIPROT:PI4728; UNIPARC:UPI000013A17E

C:Comment: X. campestris pv. vesicatoria is the causal agent of bacterial leaf spot dise

C:Genetics:

A:Start codon: GTG

#### Alignment Scores:

Pred. No.:	0.005	Length:	784
Score:	148.50	Matches:	152
Percent Similarity:	30.1%	Conservative:	58
Best Local Similarity:	21.8%	Mismatches:	238
Query Match:	3.4%	Gaps:	249
DB:	2		40

US-10-768-158-1 (1-2419) x JQ0317 (1-784)

```
QY 58 GGGAGTTCAGAGCAAGTACTTCGAGT-----TCCATGGCG 93
Db 137 GlyAsnSerArgCysThrValSerSerAlaCysLeuProProSerTrpLeuAlaMetAla 156
QY 94 TCGGGCTGCGCCCTTCTCGCGGGGAAGATGGAGGAGATCGCAACTTCCCGGTGCGGC 153
Db 157 ThrThrCysSerGlyValArgProTrpAlaTrpHisSerThrGlyAsnSerArgCys--- 175
QY 154 CCAGCCAGCTGTGGATCGTCACTACCCCAAGTCCG-----GCACGAGCTTGC 201
Db 176 -----ThrValSerSerAlaGlyLeuProProLeuLeuAlaMetAlaThrThrCys 193
QY 202 TGC-----AGAGGTGGTCT-----ACTGGTGAGCC 228
Db 194 CysGlyValArgProTrpAlaTrpHisSerThrGlyAsnSerArgCysThrValSerSer 213
QY 229 AGGGCGCTGACCCGATGAGATCGGCTTGATGAACATCGAGCAGCTCCGGTCTCG 288
Db 214 AlaCysLeuProProSerTrpLeuAlaMetAlaThrCysSerGlyValArgProTrp 233
QY 289 AGTACCACAGCGCGGCTTGGATCATCAAGAACTGACCTCCCGCGCTCATCAAGA 348
Db 234 -----AlaTrpHisSerThrGlyAsnSerArgCysThrValSerSer--- 247
QY 349 GGCACCTGCCCTACCGCTTCTGCCCTCTGACCTCCACATGAGAGACTCAAGGTCACT 408
Db 248 AlaCysLeuProProSerTrpLeuAlaMetAlaThrThrCysSerGlyValArgProTrp 267
QY 409 ATATGGCTCGCAACCCCAAGGATCTGGTGTGTCTTATTATCATGTTCCACCGCTCTGC 468
Db 268 AlaTrpHisSerThrGlyAsnSer---ArgCysThrValSerSerAla-----Cys 283
QY 469 GGACCATGAGCTACCGAGGCACCTTTCAAGAATTCTCGCGAGGTTTATGAATGATAAGC 528
Db 284 LeuPro-----ProSer 287
QY 529 TGGGCTACGGCTCTCGTTTGACACATGTCGAGAGTTCTGGGAGCAGCCGATCGACTCGA 588
Db 288 TrpLeuAlaMet-----AlaThrCysSerGlyValArgProTrpAlaTrpHisSer 305
QY 589 ACG-----TGCCTTTTCTCAAGT----- 606
Db 306 ThrGlyAsnSerArgCysThrValSerSerAlaCysLeuProProLeuLeuAlaMet 325
QY 607 ATGAAGACATGTCATCGGACCTGGTACGATGGTGGAGCAGCTGGCCAGATTCTCTGGGG 666
Db 326 AlaThrThrCysSerGlyValArgProTrpAlaTrpHisSerThrGlyAsnSer---Ala 344
QY 667 TGTCTGTGACAGGCCACCTGGAAGCCCTGAGGAGCACTGCCACCACTGGTGGACC 726
Db 345 CysThrValSerSerAlaCysLeuProPro-----LeuLeuLeuAlaMetAlaThr 361
QY 727 AGTGCTGCAACGCTGAGGCCCTCGCCGTGGCGGGGAAGAGTTGGGCTGTGGAAGGACA 786
```

```
Db 362 ThrCysSerGlyValArg-----ProTrpAlaTrpHisSerThrGlyAsnSerArgCys 379
QY 787 TCT-----TCACCGTCTCATGAATGAGAAGTTTGACTTGGTGATATAAC 831
Db 380 ThrValSerSerAlaCysLeuProProLeuLeuAlaMetAlaThrThrCysGly 399
QY 832 AGAAGA-----TGGGAAGTGTGACCTCACGTTTGTACCTTTATTTATATAACAGA 882
Db 400 ValArgProTrpAlaTrpHisSerThrGly-----AsnSerArg 412
QY 883 AACAAACAACCTGCATGCTCACAAATACCAGACAGTCTACTAGCCAAAAGTCTCTGATGCA 942
Db 413 -CysThrValSerSerAlaCysLeu-ProProSerTrpLeuAlaMetAlaThrThrCys- 431
QY 943 TTCATTTATTCCTGCTGACAACTCTGGAAGCAGCGTGTGAACACGGCGGGAAGGA 1002
Db 431 ----- 431
QY 1003 AGAGCGGCTGAGCGAGGGAGTGTGATGATTCCTCAACCGAAGCAGCTGTCTCGCCTTTA 1062
Db 432 ---SerGlyValArgProTrpAlaTrpHisSerThrGlyAsnSerArgCysThrValSerS 451
QY 1063 GAACGTGACGCTCTCCATGCTCTGATTACAAACAGTCTCCACATTGCACT----- 1112
Db 451 erAlaCys--LeuProProSerTrpLeuAlaMetAlaThrThrCysSerGlyValArgP 470
QY 1113 -----TCCATGGCTGGACCGTAAGATAAAGCTGTATATATATGCAACT 1158
Db 470 roTrpAlaTrpHisSerThrGly-----AsnSerAlaCysThrValSerSer- 485
QY 1159 AGAATGTCTGCTCTTTCACACCCGCTATTATGTTATTT-----TATAG 1200
Db 486 -AlaCysLeuPro-----ProLeuLeuAlaMetAlaThrThrCysSerGlyValAr 503
QY 1201 AGCTTTTCTTCACTGAAATCTACATAAATGTGCTAGTAAACCAATAAAGATTTCATTTCOAAG 1260
Db 503 gProTrpAlaTrpHisSerThr-----G 511
QY 1261 GGAATCAGAGGAGGACACACCGCAATGGTAGAAGATCTCAGGGTTAACTCTTTATTTT 1320
Db 511 yAsnSerAlaCysThrValSerSerAlaCys-----LeuPr 523
QY 1321 TGTAGTTTATTATCTAAGGCACAGCATTTCTCTCACTTGTGTTCTCAGATAGTGGTG 1380
Db 523 oProLeuLeuAlaMetAlaThrThr---CysSerGly-----ValArgProTrp-- 539
QY 1381 AGAACAGAGGATGAGTTGGTCTGTTGGGGGAATCTGCACACTTGTGTTTATTCTGACGA 1440
Db 540 -----AlaTrpHisSerThrGly----- 545
QY 1441 GTTCACTTCTCAGAACCTTCTGAAATGAGCAGAAATGTTCTACTAGTCTTCAAGATG 1500
Db 546 -----AsnSerAlaCy 549
QY 1501 GAGTCTCTTCTGCCAGAGACTTCCAGCGGCGCTCCAAAGGCCCAATCAGAGGAGGCC 1560
Db 549 sThrValSerSerAlaCysLeuPro-----ProLeuLeuAlaMe 563
QY 1561 GCGGAGCATGTGCTGAGGGAAGTCTGCTCCCTGGTGGAGCTGGCAGGTGGGAGTCTAATGCAG 1620
Db 563 tAlaThrThrCysSerGlyValArgProTrp--AlaTrp----- 575
QY 1621 TCAGGAGCATTTGTCATGTCAGTGGTGGAGAGTCCGCCACCAAGAGCGAGTTGCGCTCG 1680
Db 576 -----HisSerThrGlyAsnSerArgCysThrValSerSerAlaCys----- 589
QY 1681 GAATTTGAGTGAATTCACAGCTTACTTGTGTTCTGAAAGTGATAGCTACTAATGCT 1740
Db 590 -----LeuProProPro-----LeuLeuAlaMe 597
QY 1741 GGCA-----AGCAGATG 1752
```



```
Db 2833 ProGlyProProGlyProVal-----GlyProProGlyGluArg--- 2845
QY 486 CTCGGTAGCTCATGTCCGACAGAGCGGTGAATGATAATAAGACACACAGATCCT 427
Db 2846 -----GlyPheThrGlyLysAspGlyP-oThrGlyProArgGlyProProGlyPro 2862
QY 426 TGGGTTCCGAGCCATAGATGACCTTGGAGTCTCCATTGGAGGTGACAGGGCAGAA 367
Db 2863 AlaGlyAlaPro-----GlyValProGlyValAlaGlyPro 2874
QY 366 AGCGGTAGGCGAGGTGGCTCTTTGATGAGCGGAGAGGTCAAGTTCCTTGTATGATGCCA 307
Db 2875 SerGlyLysProGlyLysPro-----GlyAspArgGlyThrPro----- 2887
QY 306 GCCCGCGTGGGTACTCCAGGA----- 283
Db 2888 -----GlyThrProGlyMetLysGlyGluLysGlyAspArgGlyAspIleAla 2903
QY 282 -----CCGGAGCTGCTGCTGA-----TGTTCATCA----- 256
Db 2904 SerGlnAsnMetMetArgAlaValAlaArgGlnValCysGluGlnLeuIleAsnGlyGln 2923
QY 255 ---AGCCGA-----TCTCATCGGGGT 238
Db 2924 MetSerArgPheAsnGlnMetLeuAsnGlnIleProAsnAspTyrTyrSerAsnArgAsn 2943
QY 237 CAGCCGCTGCTGCTACCAAGTAGACCA---CCTCTGCGAGCAAGTGTGCGGAGCTTGG 181
Db 2944 Gln---ProGlyProProGlyProProGlyProProGlyAlaAla----- 2957
QY 180 GTAGGTGACGATCCACAGCTCGCTGGCGCCGACCGGA----- 142
Db 2958 -----GlyThrArgGlyGluProGlyProGlyProGlyArgProGlyPheProGly 2973
QY 141 -----AGTTGGCGATCTCTCCATCTTCCCGCGAGAAAGCGCGCAGCCGACCCAT 88
Db 2974 ProProGlyValGlnGlyProProGlyGluArgGlyMetProGlyGluLysGlyLuarG 2993
QY 87 GGAATCGAAGTACTTGC-----TCT 67
Db 2994 GlyThrGlySerGlnGlyProArgGlyLeuProGlyProProGlyProGlnGlyGluSer 3013
QY 66 CGAATC---CCCCGGGTGCTGGGGTCTCGG---CCTCGCTCCTCCGCGCAGCCGCGC 16
Db 3014 ArgThrGlyProProGlySerThrGlySerArgGlyProProGlyProProGlyArg 3032

RESULT 82
Ti3049
e:elid - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: Ti3049
R:Treisman, J.E.; Luk, A.; Rubin, G.M.; Heberlein, U.
submitted to the EMBL Data Library, March 1998
A:Reference number: Z17592
A:Accession: Ti3049
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2715 <TRE>
A:Cross-references: UNIPROT:Q8IN94; UNIPARC:UPI000007D87E; EMBL:AF053091; NID:g2981220;
C:Genetics:
A:Gene: eld
A:Cross-references: FlyBase:FBgn0003013
C:Function:
A:Description: could act as a transcription factor antagonistic to the Wg pathway
C:Keywords: DNA binding

Alignment Scores:
Pred. NO.: 0.00782 Length: 2715
Score: 147.50 Matches: 82
Percent Similarity: 36.9% Conservative: 36
Best Local Similarity: 25.6% Mismatches: 98
```

```
Query Match: 3.4% Indels: 104
DB: 2 Gaps: 16

US-10-768-158-1 (1-2419) x Ti3049 (1-2715)
QY 759 GGCACGCGGAGGCGCTCAGCGTTGCAGACTGCTCCACAGCTGTGGGAGTCT--- 703
Db 1433 GlyProGlnProProThrGlnThrValAlaGlyProProAlaGlyGlyAlaProGly 1452
QY 702 -----CCGTCAGGGCTTCAGCTGGGCTTGTTCACAGACACCC 664
Db 1453 AlaProProSerSerAlaTyrProThrGlyArgProSerGlnGlnAspTyrTyrGlnPro 1472
QY 663 CCA----- 661
Db 1473 ProProAspGlnSerProGlnProArgArgHisProAspPheIleLysAspSerGlnPro 1492
QY 660 -----GGAATCTGCCAGCTGTCTCCA----- 640
Db 1493 TyrProGlyTyrAsnAlaArgProGlnIleTyrGlyAlaTyrGlnSerGlyThrGlnGln 1512
QY 639 -----CCATCCTCAGCAGGTCCCGATGCATGCTCTTCATCTTGAGAAAAA 595
Db 1513 TyrArgProGlnTyrProSerSerProAlaProGlnAsnTrpGlyGlyAlaProProArg 1532
QY 594 GCACGTTCCGAGTCCATCGGTGCTCCAGAACTCCTGCACGCTGCTCAAAACAGGAGCCGT 535
Db 1533 GlyAlaAlaProProProGlyAlaProHisGlyProPro-IleGlnGlnProAlaGlyVa 1552
QY 534 AGCCGAGCTTATCATTAACCTCCGCGCAGAAATTTCTTGAAGGTGCTCGTAGCTCA 475
Db 1552 IAlaGlnTrpAspGlnHisArgTyrProProGln-----G1 1564
QY 474 TGGTCCG-----CAGAGAGCGGTGGAAGTATGATAATAACACCA 436
Db 1564 nGlyProProProProGlnGlnGlnGlnGlnProGlnGln--GlnGlnGlnGlnPro 1583
QY 435 CCA-----GATCCTTGGGTTGCCAGCATATAGTAGCT---TGG 397
Db 1584 ProTyrGlnGlnValAlaGlyProProGlyGlnGlnProProGlnAlaProProGlnTrp 1603
QY 396 AGTCTCCATTGTGAGGTGACAGGCGCAGAAAGCGTAGGCGAGGTGCTCTTGATAGGC 337
Db 1604 AlaGlnMetAsnProGlyGlnThrAlaGlnSerGlyIleAlaPro----- 1618
QY 336 GGGGAGAGGTGAGTCTTGTATGATGTCAGGCGCGGTGTGGGTACTCCA---GGACCG 280
Db 1619 -----ProGlySerProLeuArgProProSerGlyPro 1629
QY 279 GGAGCTGCTCGATGTTTCATCA-----AGCCGATCTCATCGGGTGCAGCGCCTGCG 226
Db 1630 GlyGlnGlnAsnArgMetProGlyMetProAlaGlnGlnGlnGlnSerGlnGlnGlnGly 1649
QY 225 TCACCAAGTAGACACACCTCTCTGCAGCAAGCTGGTCCGAGCTTGGGTAGGTGACGATCC 166
Db 1650 GlyValProGlnProProPro-GlnGlnAlaSerHisGly--GlyVal-----Pr 1665
QY 165 ACAGTCTCGTGGCGCGCACCGGGAAGTTGGGATCTCTCCTCATCTTCCGCGGCGAGAAG 106
Db 1665 oSerProGlyLeuPro-----GlnValGly-----ProGlyGlyMetVa 1678
QY 105 GCGGCGAGCGCAGCCCATG-----GAACTCGAAGTACTTGTCTCTCGAACTCCCGCC 55
Db 1678 lLysProProTyrAlaMetProProProProSerGlnGlyValGlyGlnGlnValGlyGl 1698
QY 54 GGGTGTCTGGGGTCTCGGCTCTCGCTCTCCGCCATGCGCGCGCGCGCGCGCG 5
Db 1698 nGlyProProGly---GlyMet-MetSerGlnLysProProProMetPro 1713

RESULT 83
PIHUSD
salivary proline-rich glycoprotein precursor PRB4 (large allele) [validated] - human
N:Contains: basic proline-rich protein IB-5; proline-rich peptide P-D
```



hypothetical protein T10B10.1 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T24769  
R:Sim, M.  
submitted to the EMBL Data Library, May 1996  
A:Reference number: Z19934  
A:Accession: T24769  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-428 <WIL>  
A:CROSS-references: UNIPROT:Q22369; UNIPARC:UPI000007714C; EMBL:Z72514; PIDN:CAA96674.1;  
A:Experimental source: clone T10B10  
C:Genetics:  
A:Gene: CESP:T10B10.1  
A:Map position: X  
A:Introns: 268/1

Alignment Scores:  
Pred. No.: 0.0057 Length: 428  
Score: 147.00 Matches: 81  
Percent Similarity: 32.9% Conservative: 17  
Best Local Similarity: 27.2% Mismatches: 112  
Query Match: 3.3% Indels: 88  
DB: 2 Gaps: 19

US-10-768-158-1 (1-2419) x T24769 (1-428)

QY	786	TGTCCTCCACAGCCCACTTCCCGCGCCACCGGCGAGGCCT---CAGCGTTGCAGC	730
Db	93	CysProAlaGlyAla-----ProGlyProProGlyAsnProGlyLysArgGlyAsp	109
QY	729	ACTGTCACACAGCT-----CGTGGCAGTGTCCGTGAGGCTTCCAGCTGGGCTGT	676
Db	110	GlulGlyHisProGlyAspGluGlyArgGlyAlaSerGlyIleSerLeuAlaThr	129
QY	675	CACAGGACACCCCGAGGAATC-----TGGCCAGTGTCTCCACCATCGTCACAGTCCC	622
Db	130	HisAspIleProGlyGlyCysIleLysCysProGluGlyProAlaGlyProProGlyPro	149
QY	621	GATGCATGTCTTCATACTTGAGAAAGACCGTTCAGTCCATGCGGTCTCCAGAACT	562
Db	150	Asp-----GlyAspSerGlyProGluGlyPheProGlyLeu	161
QY	561	CTGCAGCTGCTCAACACGAGGAGCGTAGCCAGCTTATCATTAACCTCCGCGCAGA	502
Db	162	GlnGlyGlnSer-----GlyProSerGlyGluAspGlyAlaProGlyGln	176
QY	501	ATTCTTGAAGGTGCTCGTAGCTCATGTCGCGCAGAGCGGTGGAACCTATAATAAG	442
Db	177	-----GluGlyAlaProGlyAspGlnGlyGlnGlyProLysGlyTyraAspGlyThr	194
QY	441	ACACCACAGAT-----	430
Db	195	AspGlyProAspGlyGlnProGlyThrThrThrPheProGlyGlnAlaGlyGlnProGly	214
QY	429	-----CCTTGGGTTGCGAGCCATATAGATACCTTGAGTCTCATTTGGAGTACG	376
Db	215	GlulProGlyTrpLeuGlyGluPro-----GlyLeuProGlyGln	227
QY	375	AGGCAGAAAGCGTAGGCGAGTGGCTCTTGATGAGCGCGGAGAGGTTCAGTCTCTTGA	316
Db	228	HisGlyGluProGlyLysAspGlyGluGlu-----GlyProGlnGlyAlaPro---	243
QY	315	TGATGTCAGGCCCGGCT-----GTGGGTACTCCAGGA-----	283
Db	244	GlyThrProGlyAsnAlaGlyHisAspAlaPheProGlyThrProGlyGlnAlaGlyLys	263
QY	282	CCGGAGCT-----GCTCGTCATGTTTATCAAGCCGATCTCATCGGGTCAG	235
Db	264	ProGlyAlaProGlyLysAspAlaAsnTyrcysProCysProGlnArgGlnAspAspArg	283
QY	234	CGCCCTGGCTCACCAAGT-----AGACCACCTCTCTGCAGCAAGCTGGTGGCGG	187

Db	284	ThrProProSerSerGlyThrSerAlaProGlnProProProGlySer-----	300
QY	186	ACTTGGGTAGTGACGATCCACAGTCGTCGCGCGCCGACCGGAGTTGGGATCTCTCT	127
Db	301	-----ThrAlaAlaProGlyThr---ArgAlaPro	309
QY	126	CCATCTTCCCGCGCAGAGCGCGCAGCCGACCCATCGAAGTACTTGCTCTCT	67
Db	310	ProAlaThrArgAlaProProAlaThrArgAlaProProAlaThr-----Thr	325
QY	66	CGAACTCCCGCGGTGTCTGGCGGTCTCGCTCGCTCT---CGCCCATGTCGCGC	16
Db	326	ArgAlaProProAlaThrThrArgProAlaProAlaSerGlnProProValArg	343

RESULT 85  
S23810  
collagen alpha 1(XVI) chain precursor - human  
N:Alternate names: procollagen alpha 1(XVI) chain  
C:Species: *Homo sapiens* (man)  
C>Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 09-Jul-2004  
C:Accession: S23810; PQ0612; S08012  
R:Pan, T.C.; Zhang, R.Z.; Mattei, M.G.; Timpi, R.; Chu, M.L.  
Proc. Natl. Acad. Sci. U.S.A. 89, 6565-6569, 1992  
A:Title: Cloning and chromosomal location of human alpha1(XVI) collagen.  
A:Reference number: S23810; MUID:92335339; PMID:1631157  
A:Accession: S23810  
A:Molecule type: mRNA  
A:Residues: 1-1603 <PAN>  
A:CROSS-references: UNIPROT:Q07092; UNIPARC:UPI0000126D33; EMBL:M92642; NID:g180757; PID  
R:Yanaguchi, N.; Kimura, S.; McBride, O.W.; Hori, H.; Yamada, Y.; Kanamori, T.; Yamakoshi  
J. Biochem. 112, 856-863, 1992  
A:Title: Molecular cloning and partial characterization of a novel collagen chain, alpha  
A:Reference number: PQ0612; MUID:93203161; PMID:1284248  
A:Accession: PQ0612  
A:Molecule type: mRNA  
A:Residues: 'GGR', 421-536, 'P', 538-1159, 'P', 1161-1162, 'P', 1164, 'P', 1166-1603 <YAM>  
A:CROSS-references: UNIPARC:UPI000016B3C2; GB:S57132; NID:g298641; PIDN:AAB25797.1; PID:  
A:Experimental source: placenta  
R:Kimura, S.  
submitted to the EMBL Data Library, April 1989  
A:Description: Partial nucleotide and amino acid sequence of a collagen-like protein fr  
A:Reference number: S08012  
A:Accession: S08012  
A:Molecule type: mRNA  
A:Residues: 403-419, 'GR', 421-536, 'P', 538-846, 'VM', <KIM>  
A:CROSS-references: UNIPARC:UPI0000073DAB; EMBL:X14963; NID:g29984; PIDN:CAA33085.1; PID:  
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit (  
ed and subsequently O-glycosylated.  
C:Genetics:  
A:Gene: GDB:COL16A1  
A:CROSS-references: GDB:134045; OMIM:120326  
A:Map position: lp34-lp34  
A:Complex: type XVI collagen may be a homotrimer, or a heterotrimer of two alpha 1(XVI)  
C:Function:  
A:Description: structural component of extracellular fibrous polymer as a minor form pro  
A:Note: may play a role in forming elastic connections at fibril surfaces  
C:Keywords: cell binding; coiled coil; extracellular matrix; glycoprotein; hydroxylysine  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-1603/Product: collagen alpha 1(XVI) chain #status predicted <MAT>  
F:22-333/Domain: amino-terminal nonhelical #status predicted <NC11>  
F:334-1577/Region: interrupted helical  
F:334-360/Domain: collagenous COL10 #status predicted <COL0>  
F:375-505/Domain: collagenous COL9 #status predicted <COL9>  
F:521-554/Domain: collagenous COL8 #status predicted <COL8>  
F:539-541/Region: cell attachment (R-G-D) motif  
F:572-630/Domain: collagenous COL7 #status predicted <COL7>  
F:652-722/Domain: collagenous COL6 #status predicted <COL6>  
F:738-875/Domain: collagenous COL5 #status predicted <COL5>  
F:887-938/Domain: collagenous COL4 #status predicted <COL4>  
F:973-987/Domain: collagenous COL3 #status predicted <COL3>  
F:1005-1007/Region: cell attachment (R-G-D) motif

F:1011-1432/Domain: collagenous COL2 #status predicted <COL2>  
F:1226-1228/Region: cell attachment (R-G-D) motif  
F:1472-1577/Domain: collagenous COL1 #status predicted <COL1>  
F:1578-1603/Domain: carboxyl-terminal nonhelical #status pred  
F:47,327/Binding site: carbohydrate (Asn) (covalent) #status

Alignment Scores:		
Pred. No.:	0.0076	1603
Score:	147.00	Matches: 86
Percent Similarity:	33.8%	Conservative: 112
Best Local Similarity:	29.4%	Mismatches: 13
Query Match:	3.3%	Indels: 82
DB:	2	Gaps: 17

US-10-768-158-1 (1-2419) x S23810 (1-1603)

QY	775	AGCCAACTCTTCCCAGGCCCAACGGGAGGCGCT---CAGCGTTGCAGCACTGGTCCACC	719
DB	1018	SerProGlyLeuProGlyProProGlyLeu-                         ProGlyLeuArgGlyGluGluGlyProPr	1037
QY	718	AGCTGGTGGCAGTCTCGTCAGGCGCTTCCAGCTGGCGCTTGTACAGGACACCCCGAG	659
DB	1037	oGlyMetArgGlySerProGly-                 ProGlySerProGly	1047
QY	658	AATCTGCCACAGCTCTCCACCATCGTCACACAGGTCCCGATGCATGTCTTCATACTTGAGA	599
DB	1047	y-----ProIleGlyProProGlyPheProGly-                 ProIleGlyProProGlyPheProGly	1056
QY	598	AAAAGACGTTGAGTCCATCGGTGCTCCAGAACTCTGCACGTGCTCAAACAGAG	539
DB	1057	-----AlaValGlySerProGlyLeuProGlyLeuGlnGlyGluArgG1                 AlaValGlySerProGlyLeuProGlyLeuGlnGlyGluArgG1	1071
QY	538	CCGTAGC-                 CCGTAGC	521
DB	1071	yLeuThrGlyLeuThrGlyAspGlyGlyGluProGlyProProGlyGlnProGlyTyPr	1091
QY	520	TTCATAAACCCTCCGGCAGAAATCTTGAAGGTGCCTCGGTAGTCTATGGTCCCGAGAG	461
DB	1091	oGlyAlaThrGlyProProGlyLeuProGlyIleLeuLysGlyGlu-                 ArgGlyTyPrTh	1109
QY	460	CGGTGGAACGTATAATAGACACACACAGATCCTTGGGTGGCGAGCCATATAGATCAC	401
DB	1109	rglySerAlaGlyGluLysGlyGluProGlyProProGlySerGlyGlyLeu-                 Pr	1127
QY	400	TTGGAGTCTCCATTGTGGAGTCCAGAGGCGAGAAAGCGGTAGGACGTGGCTTGTATG	341
DB	1127	oGlyProProGlyProAlaGlyProAArgGlyGluArgGlyProGlnGlyAsnSer-                 TCAGTTCTCTTGATG	314
QY	340	AGCGGGGAGAGG-                 GlyGluLysGlyAspGlnGlyPheGlnGlyGlnProGlyPheThrGlyProThrG1	1164
QY	313	ATGTCCAGGC-                 CCGGTGTGGGTACTCCAGGACCGGAGCTCTCGTCGATGTC	260
DB	1164	ySerProGlyPheProGlyLysValGlySerProGlyProProGlyProGlnAlaGluLy	1184
QY	259	ATCAAGCCGATCTCATCGGGTCAAGC-                 CCTGGTCAACCAAGTAGACCA---CC	209
DB	1184	sGlySerGluGlyIleArgGlyProSerGlyLeuProGlySerProGlyProProGlyPr	1204
QY	208	TCCTGCAGCAAGCTGGT-                 GCCGACTT-----G	182
DB	1204	oProGlyIleGln-GlyProAlaGlyLeuAspGlyLeuAspGlyLysAspGlyLysProG	1224
QY	181	GGGTAGTGTGACATCCACAGTCTCGCTGGGCGGACCGCGAAGTGGCGATCTCTCCATC	122
DB	1224	lyLeuArgGlyAspProGlyProAlaGlyPro-                 ProGly---LeuMetGlyProProG1	1242
QY	121	TTCCCGCGGACGAAGGCGGACCGCGCATGGAACCTCGAAGTACTTGTCTCTCGAAC	62
DB	1242	yPheLysGlyLysThrGlyHisProGlyLeuProGlyProLysGlyAspCysGlyLysPr                 Pr	1262
QY	61	T---CCCCCGGGGTCTGGGGTCTCGGCT	34



Kidney Int. 44, 1316-1321, 1993  
A>Title: Differential splicing of COL4A5 mRNA in kidney and white blood cells: a complex  
A/Reference number: 156971; MUID:94133540; PMID:8301933  
A/Accession: 156971  
A>Status: translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1258-1276 <GU01>  
A/Cross-references: UNIPARC:UPI000016B421; GB:S69168; NID:g545095; PIDN:AAC60612.1; PID:  
A/Note: kidney splice form  
A/Accession: 176598  
A>Status: translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1284-1291 'TFLGYLACLV' <GU02>  
A/Cross-references: UNIPARC:UPI000011DDFD; GB:S69169; NID:g545097; PIDN:AAC60613.1; PID:  
A/Note: frameshift mutation in patient with Alport syndrome  
R/Myers, J.C.; Jones, T.A.; Pohjola, E.R.; Kadr, A.S.; Goddard, A.D.; Sheer, D.; So  
Am. J. Hum. Genet. 46, 1024-1033, 1990  
A>Title: Molecular cloning of alpha5(IV) collagen and assignment of the gene to the regi  
A/Reference number: A35335; MUID:90252791; PMID:2339699  
A/Accession: A35335  
A>Status: nucleic acid sequence not shown  
A/Molecule type: mRNA  
A/Residues: 1448-1477 <MYE>  
A/Cross-references: UNIPARC:UPI00001738E2  
R/Nakazato, H.; Hattori, S.; Ushijima, T.; Matsuura, T.; Koitabaehi, Y.; Takada, T.; Yoe  
Kidney Int. 46, 1307-1314, 1994  
A>Title: Mutations in the COL4A5 gene in Alport syndrome: a possible mutation in primor  
A/Reference number: 156975; MUID:95156893; PMID:7853788  
A/Accession: 156975  
A>Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1595-1602 <NAK>  
A/Cross-references: UNIPARC:UPI00000004F8; GB:S75903; NID:g913882; PIDN:AAB33374.1; PID:  
A/Note: premature termination mutation from a patient with Alport syndrome; one other mu  
R/Lemnick, H.H.; Schroeder, C.H.; Brunner, H.G.; Nelen, M.R.; Zhou, J.; Tryggvason, K.;  
Genomics 17, 485-489, 1993  
A>Title: Identification of four novel mutations in the COL4A5 gene of patients with Alpo  
A/Reference number: 154188; MUID:94010948; PMID:8406498  
A/Accession: 154188  
A>Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1604-1607 'VHDAYKC' <LEW>  
A/Cross-references: UNIPARC:UPI000011F85C; GB:S65767; NID:g425563; PIDN:AADI3967.1; PID:  
A/Note: frameshift mutation from a patient with Alport syndrome; five other mutations a  
C/Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
ed and subsequently O-glycosylated.  
C/Genetics:  
A/Gene: GDB:COL4A5; ATS  
A/Cross-references: GDB:120596; OMIM:303630  
A/Map position: Xq22-Xq22  
A/Introns: 27/3; 47/3; 77/3; 92/3; 107/3; 128/3; 146/3; 155/3; 182/3; 203/3; 215/3; 229/  
/3; 799/1; 837/1; 893/1; 923/1; 1006/1; 1036/1; 1082/3; 1125/1; 1152/1; 1185/1; 1  
A/Note: the alpha 5(IV) and alpha 6(IV) chain genes are encoded on opposite strands with  
C/Complex: this minor type IV collagen is thought to form a heterotrimer of two alpha 5  
among trimer amino-terminal domains (with disulfide and desmosine cross-links), dimeric a  
er associations in the interrupted helical domain (with disulfide and desmosine cross-li  
C/function:  
A/Description: minor structural component of extracellular basement membrane  
A/Superfamily: collagen alpha 1(IV) chain  
C/Keywords: Alport syndrome; basement membrane; coiled coil; extracellular matrix; glyco  
F.1-26/Domain: signal sequence #status predicted <SIG>  
F.27-1691/Product: collagen alpha 5(IV) chain, renal splice form #status predicted <MAT1  
F.27-1264,1271-1691/Product: collagen alpha 5(IV) chain, leukocyte splice form #status P  
F.27-41/Domain: amino-terminal nonhelical, NC2 #status predicted <NC2>  
F.42-1462/Region: interrupted helical  
F.1463-1691/Domain: carboxyl-terminal nonhelical, NC1 #status predicted <NC1>  
F.1473-1573/Domain: collagen IV carboxyl-terminal repeat <CT1>  
F.1583-1687/Domain: collagen IV carboxyl-terminal repeat <CT2>  
F.29-32,38,40,124,451,481,484/Diulfide bonds: interchain #status predicted  
F.1482-1570/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F.1482-1570/1515-1573/Diulfide bonds: (or 1482-1573, 1515-1570) #status predicted  
F.1527-1533,1638-1644/Diulfide bonds: #status predicted  
F.1592-1684,1626-1687/Diulfide bonds: (or 1592-1687, 1626-1684) #status predicted

Alignment Scores:  
Pred. No.: 0.00769 Length: 1691  
Score: 147.00 Matches: 75  
Percent Similarity: 35.8% Conservative: 22  
Best Local Similarity: 27.7% Mismatches: 108  
Query Match: 3.3% Indels: 66  
DB: 1 Gaps: 16

US-10-768-158-1 (1-2419) x S22917 (1-1691)

QY 765 TTCCCGCGCCACCGAGCGGCTTCAGCGTTGCAGCACTGGTCCACGAGCTGGTGGCAGT 706  
|||||  
DB 401 PheProGlyGluAsgGly-----GlnysgLyAspGluGlyProPro---GlyIleSer 417  
QY 705 GCTCCG-----TCAGGGCTTCACCTGGCGCTTGTTCACAGACACCCCA 661  
|||||  
DB 418 IleProGlyProProGlyLeuAspGlyGlnProGlyAlaProGlyLeuProGlyProPro 437  
QY 660 GGAATCTGGCCAGCTGCTCCACCATCGTCACCGAGTCCCGATGCTGCTTCATATTGA 601  
|||||  
DB 438 Gly-----ProAlaGlyProHisIleProProSerAspGluIleCys----- 451  
QY 600 GAAAAAGCACCTTCGAGTCCATCGGCTGCTCCAGAACTCTCGCACGCTGCTCAACACGAG 541  
|||||  
DB 452 -----GluProGlyProProGlyProProGlySerProGlyAspLysGlyLeuGln 468  
QY 540 AGCCGTAGCCAGCTTATCATTTCAATAAACCTCCCGGAGAAATCTTTGAAAGTGCTCGGT 481  
DB 469 GlyGluGlnGlyValLysGlyAsp-----LysGlyAspThrCysPheAsnCysIleGly 486  
QY 480 AGCTCATGTCGCGAGAGAGCGGTGGAAGTAAATGAACACACACAGATGCTTGGGGT 421  
DB 487 ThrGlyIleSerGlyProProGlyGlnProGlyLeuProGlyLeuProGlyPro----- 504  
QY 420 TCGAGCCATATAGATGACCTTGGAGTCTCCATTTGTGGAGGTTCAGAGGCGAGAAAGCGGT 361  
DB 505 -----ProGlySerLeuGlyPheProGlyGlnLysGlyGluLysGly 518  
QY 360 AGGCAGGTGCTCTTGATAGAGCGGGAGAGGTGCTGCTTGTGATGATGCTCCAGGCGCG 301  
DB 519 GlnAlaGlyAlaThr-----GlyProLysGlyLeuProPro---GlyIleProGly--- 533  
QY 300 GCTGTGGGTACTCCAGGA-----CCGGAGGTGCTCGTCGATGT 262  
DB 534 AlaProGlyAlaProGlyPheProGlySerLysGlyGluProGlyAspIleLeuThrPhe 553  
QY 261 TCATCAAGCCGATCTCATCGGGGTGAG-----CGCCCTGGCTCACCAAGTAGACCA--- 211  
DB 554 ProGlyMetLysGlyAspLysGlyGluLeuGlySerProGlyAlaProGlyLeuProGly 573  
QY 210 -----CCTCTGAGCAAGCTGGTGGCGGACTGGGTAGGTAGGTACACACGT 160  
DB 574 LeuProGlyThrProGlyGlnAspGlyLeu-----ProGlyLeu 586  
QY 159 CGCTGGGCGGCACCGGGAGTTGGCGATCTCTCCATCTTCCGCGCGCAGAGGCGGCA 100  
DB 586 uProGlyProLysGlyGluProGlyGlyLe-ThrPhe-----LysGlyGluA 602  
QY 99 GCCGACGCCCTTGAACCTCGAAGTACTTGTCTCGAAGTCCCGCGGTGCTGGGGTCT 40  
DB 602 rgGlyProProGlyAsn-----ProGlyLeuProGlyLeuP 614  
QY 39 CGGCTCGCTCTCCGCCATGCGCGCGCG 11  
DB 614 roGlyAsnIleGlyProMetGlyProPro 623

RESULT 87

CG801S  
collagen alpha 1(I) chain - bovine (tentative sequence) (fragments)  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 24-Apr-1984 #sequence revision 31-Dec-1993 #text change 09-Jul-2004  
C/Accession: A91193; A91229; A91387; A91201; A91200; A43048; A02853



R;Rauterberg, J.; Timpl, R.; Furthmayr, H.  
 Eur. J. Biochem. 27, 231-237, 1972  
 A:Title: Structural characterization of N-terminal antigenic determinants in calf and hu  
 A:Reference number: A91193; MUID:72255334; PMID:4115172  
 A:Accession: A91193  
 A:Molecule type: protein  
 A:Residues: 1-19 <RAU>  
 A:CROSS-references: UNIPROT:P02453; UNIPARC:UPI0000173B51  
 A:Experimental source: skin  
 A:Note: the epsilon carbon of Lys-9, by homology with the rat alpha 1(I) chain, is conve  
 R;Fietzek, P.P.; Kuehn, K.  
 Eur. J. Biochem. 52, 77-82, 1975  
 A:Title: The covalent structure of collagen: amino-acid sequence of the cyanogen-bromide  
 A:Reference number: A91229; MUID:76022320; PMID:1164916  
 A:Accession: A91229  
 A:Molecule type: protein  
 A:Residues: 20-145 <FIE>  
 A:CROSS-references: UNIPARC:UPI0000173B52  
 A:Experimental source: skin  
 A:Note: Lys-103 is hydroxylated and binds glucosylgalactose  
 R;Fietzek, P.P.; Wendt, P.; Keil, I.; Kuehn, K.  
 FEBS Lett. 26, 74-76, 1972  
 A:Title: The covalent structure of collagen: amino acid sequence of alpha1-CB3 from calf  
 A:Reference number: A91387; MUID:73049499; PMID:4673951  
 A:Accession: A91387  
 A:Molecule type: protein  
 A:Residues: 146-294 <F12>  
 A:CROSS-references: UNIPARC:UPI0000173B53  
 A:Experimental source: skin  
 R;Fietzek, P.P.; Rexrodt, F.W.; Hopper, K.E.; Kuehn, K.  
 Eur. J. Biochem. 38, 396-400, 1973  
 A:Title: The covalent structure of collagen. 2. The amino-acid sequence of alpha1-CB7 fr  
 A:Reference number: A91211; MUID:74086118; PMID:4359390  
 A:Accession: A91211  
 A:Molecule type: protein  
 A:Residues: 295-562 <F13>  
 A:CROSS-references: UNIPARC:UPI0000173B54  
 A:Experimental source: skin  
 R;Wendt, P.; Mark, K.V.D.; Rexrodt, F.; Kuehn, K.  
 Eur. J. Biochem. 30, 169-183, 1972  
 A:Title: The covalent structure of collagen. The amino-acid sequence of the 112 residues  
 A:Reference number: A91201; MUID:73042276; PMID:4343808  
 A:Accession: A91201  
 A:Molecule type: protein  
 A:Residues: 563-675 <WEN>  
 A:CROSS-references: UNIPARC:UPI0000173B55  
 A:Experimental source: skin  
 R;Fietzek, P.P.; Rexrodt, F.W.; Wendt, P.; Stark, M.; Kuehn, K.  
 Eur. J. Biochem. 30, 163-168, 1972  
 A:Title: The covalent structure of collagen. Amino acid sequence of peptide alpha1-CB6-C  
 A:Reference number: A91200; MUID:73042275; PMID:4343807  
 A:Accession: A91200  
 A:Molecule type: protein  
 A:Residues: 676-758 <F14>  
 A:CROSS-references: UNIPARC:UPI0000173B56  
 A:Experimental source: skin  
 A:Note: Pro-726 is the only 3-hydroxyproline and the only hydroxylated proline in positi  
 R;Rauterberg, J.; Fietzek, P.; Rexrodt, F.; Becker, U.; Stark, M.; Kuehn, K.  
 FEBS Lett. 21, 75-79, 1972  
 A:Title: The amino acid sequence of the carboxyterminal nonhelical cross link region of  
 A:Reference number: A43048  
 A:Accession: A43048  
 A:Molecule type: protein  
 A:Residues: 759-779 <RA2>  
 A:CROSS-references: UNIPARC:UPI0000173B57  
 A:Experimental source: skin  
 C:Comment: Lysines at positions 115, 124, 274, 346, 424, 496, 658, and 670 may be hydrox  
 C:Comment: Prolines in the third position of the tripeptide repeating unit (G-X-Y) are h  
 C:Comment: The order of the eight CNBR peptides in the alpha 1(I) chain of bovine skin d  
 9, 149, 268, and 217 residues.  
 C:Comment: The complete chain contains 1052 residues.  
 C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;  
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; pyroglutamic acid; trimer;

F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

#### Alignment Scores:

Pred. No.: 0.00709 Length: 779  
 Score: 146.50 Matches: 83  
 Percent Similarity: 29.9% Conservative: 13  
 Best Local Similarity: 25.9% Mismatches: 112  
 Query Match: 3.3% Indels: 113  
 DB: 1 Gaps: 19

US-10-768-158-1 (1-2419) x CGB01S (1-779)

QY	783	CCTTCACAGCCCAACTCTTCCCGGCCCA-----CGGGCAGGCGCTCAGGCTTCACGC	730
DB	432	ProProGlyAlaThrGlyPheProGlyAlaAlaGlyArgValGlyProPro-----	448
QY	729	ACTGGTCCACACAGCTGGTGGCAGTGTCTCGTCAGGCTTCACAGCT-----	685
DB	449	-----GlyProSerGlyAsnAlaGlyProProGlyProProGlyProAlaGlyLysGlu	466
QY	684	-----GGCCTTGTCCACAGGACCCCGCAGGAATCTGCCA-----GCTGCTCCA	640
DB	467	GlySerLysGlyProArgGlyGluThrGlyProAlaGlyArgProGlyGluValGlyPro	486
QY	639	CCATCGTCACAGCTCCGATGCATGTCTTCACTTGAGAAAA-----GCACGT	589
DB	487	ProGlyProProGlyProAlaGly-----GluLysGlyAlaProGlyAla	501
QY	588	TCGAGTCCATGCGGTCTCCAGAACTCTGTCACGTCTCAACACGAGGCGGTAGCCCA	529
DB	502	AspGlyProAlaGlyAlaProGlyThrProGlyProGlnGlyIleAlaGlyGlnArgGly	521
QY	528	GCTTATCATTTCAAAACCTCCCGCAGAAATCTTGAAGGTGCTC-----GGT	481
DB	522	ValValGly-----LeuProGlyGlnArgGlyGluArgGlyPheProGlyLeuProGly	539
QY	480	ACGTCATGG-----TCCGAGAGACGGTGGAACTGTAATAAGACACACCA	433
DB	540	ProSerGlyGluProGlyLysGlnGlyProSerGlyAlaSerGlyLysGlyProPro	559
QY	432	GATCTTGGGTTGCGAGCCATATAGATGACCTTCGAGTCTCCATTGTGAGGTACAGG	373
DB	560	GlyProMetGly-----ProProGlyLeu-----AlaGlyProPro	571
QY	372	GCAGAAAGCGTAGGCGAGGTGGCTCTTGTATGAGCGGGAGAGGTCACTTCTCTGATGA	313
DB	572	GlyGluSerGlyArgGly-----	578
QY	312	TGTCAGGCGCGGTGGTGGTACTCCAGGA-----CCGGAGCTGCTGCTCGA	265
DB	579	AlaProGly--AlaGluGlySerProGlyArgAspGlySerProGlyAla-----	594
QY	264	TGTTTCATCAAGCCGATCTCATCGGGGTGACGCG-----CCT	229
DB	595	-----LysGlyAspArgGlyGluThrGlyProAlaGlyAlaProGlyProPro	610
QY	228	GGCTCACCAGTAGACCACTCTCTCCAGCAAGCTGGTCCGACACTGGGTAGTGACGA	169
DB	611	GlyAlaProGlyAlaProGlyProValGlyProAlaGlyLysSer-----	625
QY	168	TCCACACGTGCTGGCGCGCAGGAGTGGCGCATCTCTCCATCTTCCCGCGGCAGA	109
DB	626	GlyAspArgGlyGluThrGlyProAlaGlyProLysGlyProValGlyProAlaGlyAla	645
QY	108	AGGGCGGCGCAGCGCCATCGAAGTACTTGTCTCGAACT-----	61
DB	646	ArgGlyProAlaGlyProGlnGlyProArgGly***LysGly***ThrGly***Gly	665
QY	60	-----CCCCCG	55
DB	666	***ArgGlyIleLysGlyHisArgGlyPheSerGlyLeuGlnGlyProProGlyProPro	685
QY	54	GGGTCTGGGGTCTCGG-----CCTGCTCTCGG	25

Db 686 GlySerProGlyGluGlnGlyProSerGlyAlaSerGlyProAlaGlyProArgGlyPro 705  
QY 24 CCA 22  
Db 706 Pro 706  
RESULT 88  
T35878  
hypothetical protein SC9B10.09 - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T35878  
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, November 1997  
A:Reference number: Z21592  
A:Accession: T35878  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1039 <OLI>  
A:Cross-references: UNIPROT:O50516; UNIPARC:UPI00000DAB7A; EMBL:AL009204; PIDN:CAA15799.  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SC9B10.09

Alignment Scores:  
Pred. No.: 0.00824 Length: 1039  
Score: 146.00 Matches: 78  
Percent Similarity: 35.3% Conservative: 31  
Best Local Similarity: 25.2% Mismatches: 121  
Query Match: 3.3% Indels: 79  
Db: 2 Gaps: 16

US-10-768-158-1 (1-2419) x T35878 (1-1039)

QY 840 CCATCTCTGTTTATACACCAAGTCAAACCTCTCATTCATGGAGAGGTGCTCT 781  
Db 673 ProAlaSerGlyAlaThrProGlyThrAlaProHisAla-----AlaPro 687  
QY 780 TCCACAGCCCAACTCTTCCCGGCCCCACGGGAGGCGCTCAGCGTTCAGCAGCACTGGTCCA 721  
Db 688 GlyProAlaAspGlnLeuProSerGlyAlaSerGlyPro-----SerGlyPro 703  
QY 720 -----CCAGCTGGTGGCAGTCTCGTCAGCGGCTCCAGCTGGGCGCTTGTCCAGGACA 667  
Db 704 SerGlyProSerGlyProSerGlyProSerGlyProSerGlyProSerGlyProSer 722  
QY 666 CCCCAGGAATCTGGCCAGCTCTCCACCATCGTCACAGGTCCCGAGTCCATGCTCTTCAT 607  
Db 723 GlyProSer---GlyProSerGlyProSerGlyProProArgProSerAla----- 738  
QY 606 ACTTGAGAAAAGCAGTTCGAGTCCATGCGGTGCTCCAGAACTCTCGACGTGCTCAA 547  
Db 739 -----AlaProGlyArgProProGlyValValArgAlaProGluGlyAlaArg 754  
QY 546 ACCAGAGCGGTAGCCAGCTTATCATTAACCTCCGAGCAATCTTCAAGAGTGC 487  
Db 755 Leu-----IleProAlaTy-----ProAlaAlaGluArgAlaValArgAla 768  
QY 486 CTCGGT-----AGCTCATGTCGCGCAGAGAGCGGTGAAGTATAATAAG 442  
Db 769 LeuGlyGluAlaValAlaTyAlaGlnTrpArgArgGluAlaAla----- 783  
QY 441 ACACACCAGATCTTTGGGGTTGCGAGCCATATAGATGACCTTGGAGTCTCCATTGTGGA 382  
Db 784 -----AspPro-GlyLysValProGluTyArgGluAspIle----- 794  
QY 381 GGTACAGGGCGAAGACGGTAGGCGAGTGGCTCTTGATAGGCGGGGAGAGGTGAGTT 322  
Db 795 -----AspGluLysGlyAlaAlaGlnGlnIleAlaGlyHisLeuAlaArgGlyGlnG 812  
QY 321 CCTGTATGATGTCCAGGCCCGCGTGTGGGTACTTCCAGGACCGGAGCTGCTC----- 270

Db 812 yLeu-----ThrLeuGlySerGlnGlnThrCysAspLeuLeuGlyArgTy 827  
QY 269 ----GTCGATGTTTCATCAAGCC-----GATCTC----- 246  
Db 827 rGlyValHisValHisArgAlaLeuProAlaAlaThrProAspAspAlaAlaAlaAl 847  
QY 245 -----ATCGGGGTCAGCGCTCGCTCACCAGGTAGACCACCTCTCT 205  
Db 847 aArgThrIleGlyTyProValAlaLeuLeuAlaThrAlaProHisLeuArgHisArgAl 867  
QY 204 GCAGCAAGCTGTCGCGGACTTGGGTAGGTACGATCCACACGTCGTCGGCGCCACCG 145  
Db 867 aspLeuGlyGlyValArgLeuAspLeuAlaAspGluGlnLeuArgAlaTyAl 887  
QY 144 GGAAGTTGGCAGTCTCTCCATCTCTCCGCG-----GCAGAGGGCGGCA 100  
Db 887 aGluLeuThrGluLeuPheGlyLysProAlaGluLeuArgProValValGlnGlyMetAl 907  
QY 99 GCGCAGCCCATCGAAGTCTGGAAGTACTTGTCTCGAATCCCGGGGTGCTGGGGGTCT 40  
Db 907 aProArgGlyValAspThrValValArgAlaValIleAspProAlaAlaGlyAlaVal 927  
QY 39 CGGCCTCGCTCTCGGCCCATGCGGCC 15  
Db 927 uSerPheGlyLeuAlaGlyAlaAla 935  
RESULT 89  
PIHUPF  
N:Alternate names: basic proline-rich precursor PRB2 [validated] - human (fragment)  
N:Contains: basic proline-rich peptide IB-4; basic proline-rich peptide P-F  
C:Species: Homo sapiens (man)  
C:Date: 15-Nov-1984 #sequence\_revision 12-Apr-1996 #text\_change 05-Oct-2004  
R:Maeda, N.; Kim, H.S.; Azen, E.A.; Smithies, O.  
J. Biol. Chem. 260, 11123-11130, 1985  
A:Title: Differential RNA splicing and post-translational cleavages in the human salivary  
A:Reference number: A92492; MUID:85289325; PMID:2993301  
A:Accession: E25372  
A:Molecule type: mRNA  
A:Residues: 1-251 <MAE>  
A:Cross-references: UNIPROT:P02812; UNIPARC:UPI0000132338; GB:K03208; NID:gi90509; PIDN:  
R:Mamula, P.W.; Morley, D.J.; Larsen, S.H.; Karn, R.C.  
Biochem. Genet. 26, 165-175, 1988  
A:Title: Expression of human salivary protein genes.  
A:Reference number: A60827; MUID:88240287; PMID:3288192  
A:Accession: A60827  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 175-251 <VAM>  
A:Cross-references: UNIPARC:UPI0000174513  
R:Saitoh, E.; Isemura, S.; Sanada, K.  
J. Biochem. 93, 883-888, 1983  
A:Title: Complete amino acid sequence of a basic proline-rich peptide, P-F, from human p  
A:Reference number: A03294; MUID:83265674; PMID:6874669  
A:Accession: A03294  
A:Molecule type: protein  
A:Residues: 134-194 <SAL>  
A:Cross-references: UNIPARC:UPI0000174514  
A:Experimental source: saliva  
R:Kauffman, D.L.; Bennick, A.; Blum, M.; Keller, P.J.  
Biochemistry 30, 3351-3356, 1991  
A:Title: Basic proline-rich proteins from human parotid saliva: relationships of the cov  
A:Reference number: A38355; MUID:91190884; PMID:1849422  
A:Accession: B38355  
A:Molecule type: protein  
A:Residues: 134-194 <KAU>  
A:Cross-references: UNIPARC:UPI0000174514  
A:Experimental source: saliva  
A:Note: this peptide, which is closely related to that of peptide P-E, contains three 21  
A:Accession: A38355  
A:Molecule type: protein  
A:Residues: 10-67,'R' <KA2>





```

Db      297 GlyPheProGlyAlaProGlyAlaLysGlyGluAlaGlyProThrGlyAlaArgGlyPro 316
QY      261 TCATCAAGCCGATCTCATCGGGTCAGCGCCCTGGCTCACCAGTAGACCACTCTCTGCA 202
Db      317 GluGlyAlaGlnGlyProArgGlyGlu---ProGlyThrProGlySerProGlyProAla 335
QY      201 GCAAGCTGTGCGGGACTTGGGGTAGGTGACGATCCACACGTCGTCGGCCGACCGGGA 142
Db      336 Gly-AlaAlaGlyAsnProGlyThrAspGlyIlePro-GlyAlaLysGlySerAlaGlyA 355
QY      141 AGTGGCGATCTCTCCATC-----TTCGCGCGGAGAGGCG-----G 103
Db      355 laProGlyIleAlaGlyAlaProGlyPheProGlyProArgGlyProProGlyProGlnG 375
QY      102 GCAGCGCACGCGCATCGAACTCGAAGTACTTGTCTCGAACTCCCGCGGGGTCTGGGG 43
Db      375 lyAlaThrGlyProLeuGlyProLysGly--GlnThrGlyGluProGlyIleAlaGlyP 394
QY      42 TCTCGGCTCGCTCTCCGCATCGCCGCGCGTGGCGGTGCG 2
Db      394 heLysGlyGluGlnGlyProLysGlyGluProGlyProAla 407

RESULT 92
T19732
hypothetical protein C34F6.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T19732
R:White, S.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19171
A:Accession: T19732
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-283 <WIL>
A:Cross-references: UNIPROT:O17642; UNIPARC:UPI0000082853; EMBL:Z81479; PIDN:CAB03942.1;
A:Experimental source: clone C34F6
C:Genetics:
A:Gene: CESP:C34F6.3
A:Map position: X
A:Introns: 64/2

Alignment Scores:
Pred. No.: 0.00807 Length: 283
Score: 144.50 Matches: 78
Percent Similarity: 29.4% Conservative: 19
Best Local Similarity: 23.6% Mismatches: 93
Query Match: 3.3% Indels: 140
DB: 2 Gaps: 17

US-10-768-158-1 (1-2419) x T19732 (1-283)

QY      841 CCCACTCTCTTTATACACCAAGTCAAACTCTCATTCATGGAGACG----- 794
Db      36 ProLeuPheCysValHisMetGlnSerValThrSerGlyLeuSerGluGluLeuLeuPhe 55
QY      793 -----GTGAAGATGTCCTTCCACAGCCCAACTCTCCCGGCGCC 755
Db      56 CysLysSerLysAsnValTyrIleLysGlyGluIleGluLysSerValThrArgGlu 75
QY      754 ACGGGCAGGCGCTCAGCGT-----TGCAGCACTGGTCCACCACTGG 713
Db      76 AlaGlyArg-GlnLysArgGlnThrProGlnThrCysSerCysGly-----IleG 93
QY      712 TGGCACTGCTCGCTCAGGCTTCCAGCTGGCTTCTCACAGACACCCCGAGGAATCTG 653
Db      93 yGluThrGlyProAlaGlyValProGlyGln-----GluGlyAlaProGlyAsnAspG 111
QY      652 GCCAGCTGCTCCACCACTGTCACCAAGTCCCGATGCA----- 616
Db      111 yLysAlaGlyGlnProGlyAlaAspAlaAspGluGlnGlyPheHisTyrLy 131
QY      615 -----TGCTTCATCTTGAGAAAAAGCACGTTCCAGTCCATGGGTGCTCC 569

```

```

Db      131 sAlaProGluPheCysPhe-----AspCysProAlaGlyProPr 144
QY      568 CAGAACTCTCTGCACGTCTCAAAACGAGGAGCGTAGCCAGCTTATCATTAACCTC 509
Db      144 o----- 144
QY      508 CGCGAGAATCTTTGAAAGTGCTCGGTAGCTCATGGTCCGAGAGAGCGGTGGAACCTGA 449
Db      145 -GlyAlaVal-----GlyGlyProG 151
QY      448 TAATAAGACACACACAGATCTTTGGGTTGCGAGCCATATAGATGACCTTGGAGTCTCCA 389
Db      151 yProLysGlyProProGlyProProGly----- 160
QY      388 TTGTGGAGTCCAGAGGCGAGAAAGCGTAGGCGAGGTGGCTCTTGATGAGGCGGAGAG 329
Db      161 -----GlyProGlyGluLeuGlyGlyProGlyArgGlyGlyAsnAr 174
QY      328 GTCACTCTCTTGATGATGTCCAGGCCCGGCTGTGGGTACTCCAGGA----- 283
Db      174 ggLyPro-----ProGlyProArg---GlyProProGlyGluAlaGlyProAs 189
QY      282 -----CGGGAGCTGCTCGTCGATGTTTCATCAAGCCGATCTCATCGGG 239
Db      189 pGlyGluGlyArgProGlyGlnAlaGlyGlnThrArgSerAlaProSerProG 209
QY      238 TCAG-----CGCCCTGGCTCACCAAGTAGACCACTCTCTCGACCAAGCTG 194
Db      209 yGlnProGlyGlnProGlyGluProGlySerProGlyGluProGlyProAsp----- 226
QY      193 GTCCGCGACTTGGGTAGGTGACATCCACACGTCGCTGGGCGCCGCGGAAAGTTGGCG 134
Db      227 -----GlyArgAlaGlyHisProGlyArgAsnGlyProProGly----- 239
QY      133 ATCTCTCTCATCTTCCCG-----GGCAG 110
Db      240 ----ProProGlyAspAsnGlyGlyGlnGlyGluProGlyLysAspGlyGluAspGly 258
QY      109 AAGGCGCGCAGCGCCGACCATGGAATCGAAGTACTTCTCTCGAACTCCCCCGGGGTG 50
Db      258 uAsnGlyAlaAlaGlyAlaAlaGlyProLysGlySerCysAspHisCysProPro----- 276
QY      49 CTGGGGGTCTCGGCTCGCTCTCTCCGCCA 22
Db      277 -----ProArgThrAlaPro 281

RESULT 93
S34665
collagen, cuticular - root-knot nematode (Meloidogyne incognita)
C:Species: Meloidogyne incognita
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S34665
R:van der Eycken, W.V.; de Almeida Engler, J.; van Montagu, M.; Gheysen, G.
submitted to the EMBL Data Library, July 1993
A:Description: Identification and analysis of a cuticular collagen gene from the plant-
A:Reference number: S34665
A:Accession: S34665
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-286 <VAN>
A:Cross-references: UNIPROT:Q25467; UNIPARC:UPI000016BF3F; EMBL:Z24734; NID:g395144; PFI
Alignment Scores:
Pred. No.: 0.00809 Length: 286
Score: 144.50 Matches: 64
Percent Similarity: 34.2% Conservative: 17
Best Local Similarity: 27.0% Mismatches: 97
Query Match: 3.3% Indels: 59
DB: 2 Gaps: 11

US-10-768-158-1 (1-2419) x S34665 (1-286)

```

```

QY 735 TGCAGCAGCTGGTCCACAGCTGGTGGCAGTGTCTCCGTCAGGCTTCCAGCTGGGCTTGT 676
Db 80 CysAlaProGlyProPro-----GlyProProGlyProProGlyGlnProGly 95

QY 675 CACGAGCACCCCGAGGAATCGCCAGCTGTCTCCACCATCTGTCCACAGGTCCTCCGATGCA 616
Db 96 HisProGlyGlnProGly-----HisValGlyGlnProGlySerProGlyGlnProAla 113

QY 615 -----TGCTCTCATCTTGAGAAAAGCAGCTTCAGTCCATGGTGGTCTCCAGAACT 562
Db 114 ProProCysProLeuProGlnGlnAlaCysGlnArgCysProAlaGlyAlaProGlyThr 133

QY 561 CCGTCAGCTGTCTCABACAGGAGCGTAGCCAGCTTATCATTTATAAACCTCCGSCAGA 502
Db 134 ProGlyLys-----GlnGlyProAlaGlyGlnProGlyGlnProGlyArg 148

QY 501 ATTCTTAAAGTGCTCGTAGCTCATGTGTCGCGCAGAGCGGTGGAACCTATAATAAG 442
Db 149 -----ProGlyAlaProGlyLysSerSerGlyAla 158

QY 441 ACACACCATGATCTTGGGTTGTCGAGCCATATAGATGACCTTGGAGTCTCCATTGTGGA 382
Db 159 GlyProProGlyProAlaGlyProGlnGlyProProGlyProAlaGlyLysHisGlyGly 178

QY 381 ---GTCAGAGGCGAGAAAGCGTAGGCGAGTGGCTCTTGATGAGCGGGGAGAGTCA 325
Db 179 ProGlyGlnProGlyGlnProGlyLysAenGlyValSer-----HisProThr 194

QY 324 GTTCTTGATGATGTCCAGGCCGCTGTGGTACTTCCAGGACCGGAGCTCTCTGTCGA 265
Db 195 IlePro-----GlyProLysGlyProSerGlySer----- 204

QY 264 TGTTCATCAAGCCGATCTCATCGGGTTCAGCGCCCTGGCTCACCAAGTAGACCACTCT 205
Db 205 -----ProGlyGlnProGlyLysProGlyPro 213

QY 204 GCAGCAAGCTGGTCCGCGACTTGG-----GGTAGGTGACGA 169
Db 214 AlaGlyValAlaGlyLysThrGlyProGluGlyProProGlyProValGlyProAlaGly 233

QY 168 TCACAGCTCGCTGGCGCCGACCGGGA-----AGTTGGCGATCTCTCCATCTTCCCGC 115
Db 234 ProSerGlyLysProGlyAlaProGlyGlnProGlyProHisGlyProProGlyGlnPro 253

QY 114 GCACAGAGCGCGGACCGCCAGCCATGGAACCTCAAGTACTTGTCTCTGA 64
Db 254 GlyGlnAspAlaGlnTyrCysProCysProProArgSerLeuCysSerArg 270

RESULT 94
A:1219
collagen 1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #text_change 09-Jul-2004
C:Accession: A31219; T37290; T37291
R:Kramer, J.M.; Cox, G.N.; Hirsh, D.
Cell 30, 599-606, 1982
A:Title: Comparisons of the complete sequences of two collagen genes from Caenorhabditis
A:Reference number: A30826; MUID:83050944; PMID:7139711
A:Accession: A31219
A:Molecule type: DNA
A:Residues: 1-296 <KRA>
A:Cross-references: UNIPROT:P08124; UNIPARC:UPI0000127203; GB:V00147; NID:G6677; PIDN:CA
A:Accession: T37290
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-296 <KR2>
A:Cross-references: UNIPARC:UPI0000127203; EMBL:J01047; PIDN:AAA27988.1
A:Accession: T37291
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-296 <KR3>
A:Cross-references: UNIPARC:UPI0000127203; EMBL:V00147; NID:G6677; PIDN:CAA23463.1; PID:
C:Genetics:

```

```

A:Gene: col-1
A:Introns: 54/3; 280/2

Alignment Scores:
Score: 0.00815 Length: 296
Pred. No.: 144.50 Matches: 68
Percent Similarity: 28.4% Conservative: 10
Best Local Similarity: 24.7% Mismatches: 92
Query Match: 3.3% Indels: 105
DB: 2 Gaps: 13

US-10-768-158-1 (1-2419) x A31219 (1-296)

QY 780 TCCACAGCCCAACTTCTCCCGCCACGCGGCGCTCAGCGT----- 736
Db 75 ThrThrArgGlnAlaTyrGlyProGluValAsnProAlaProAsnLeuGlnCysGlu 94

QY 735 -----TCGACACTGGTCCACCACTGGTGGCAGTGTCTCCGTCA----- 697
Db 95 GlyCysCysLeuProGlyProProGlyProAlaGlyAlaProGlyLysProGlyLysPro 114

QY 696 GGGCTTCCAGCTGGGCTTGTTCACAGGACACCCCGGAAATCTGGCCAGCTGTCTCA--- 640
Db 115 GlyArgProGlyAlaProGlyThrProGlyLysProProValAlaProCys 134

QY 639 ---CCATCGTCACAGGTCCTGATGTCATCTTATCTTACTTGAAGAAAGCAGCTCGAGT 583
Db 135 GluProThrThrProProProCysLysProCysProGlnGlyProProGlyProProGly 154

QY 582 CCATGCGGTGTCTCCAGAACTCTTCAGCTGTCTCAACACGAGGACCGTAGCCAGCTTAT 523
Db 155 ProProGlyAlaProGlyAspProGlyGluAlaGlyThrProGlyArg----- 170

QY 522 CATTCAATAAACCCTCGGCGAGAATTCTTGAAGGTGCTCGGTAGTCTCATGTGTCGACAG 463
Db 171 -----ProGlyThrAspAlaAla 176

QY 462 AGCGGTGAACCTGATAATAAGACACCACCATCTCTGGGGTTGCGAGCCATATAGATGA 403
Db 177 ProGlySerProGlyProArgGlyProProGlyPro----- 188

QY 402 CTTGGAGTCTCCATTGTGGAGGTGAGAGGCGAGAAAGCGGTAGGCGAGGTGGCTCTTGA 343
Db 189 -----AlaGlyGluAlaGly----- 193

QY 342 TGAGCGGGGAGAGTCAAGTTCCTTGATGATGTCAGGCCCGCGCTGTGGTACTCCAGGA 283
Db 194 -----AlaProGlyProAla---GlyGluProGly 202

QY 282 CCGGAGCTGTCTGTCGATGTTTCATCAAGCCGATCTCATCGGGTTCAGGCGCTCGCTCA 223
Db 203 ThrProAlaIleSerGluProLeuThr-----ProGlyAla 214

QY 222 CCAAGTAGACCA-----CCTCTGCGACGCAAGCTGGTGGCGGACT 184
Db 215 ProGlyGluProGlyAspSerGlyProProGlyProProGlyProPro----- 230

QY 183 TGGGAGTAGTGACGATCCACACGTCGCTGGGCCCGCCAGCGGAAGTTGGCGATCTCTCCA 124
Db 231 -----GlyAlaProGlyAsn---AspGlyProPro 239

QY 123 TCTTCCCGCGCAGAAAGCGCGCAGCCGACGCGCATCGAACTCGAAGTACTTGCTCTCGA 64
Db 240 GlyProProGlyProLysGlyAlaProGlyProAspGly----- 252

QY 63 ACTCCCGCGGGTGTGGGGG---TCTCGGCTCGCTCTCGGCCA 22
Db 253 ---ProProGlyAlaAspGlyGlnSerGlyProProGlyProPro 266

RESULT 95
CGHUC6
N:Alternate names: procollagen alpha 1(II) chain
collagen alpha 1(II) chain precursor [validated] - human

```

N;Contains: chondrocalcin; collagen alpha 1(II) chain precursor splice form 1; collagen  
 C:Species: Homo sapiens (man)  
 C:Date: 28-May-1986 #sequence revision 01-Sep-1995 #text change 31-Dec-2004  
 C:Accession: A38513; S06715; S24270; A24828; S06496; A35428; A30147; A33116; S64674; S63750; I37251; I37252; I37253; I37254; I55338; I59535; I61910  
 R:Ryan, M.C.; Sieraski, M.; Sandell, L.O.  
 Genomics 8, 41-48, 1990  
 A:Title: The human type II procollagen gene: identification of an additional protein-coding exon  
 A:Reference number: A38513; MUID:91184811; PMID:2081599  
 A:Accession: A38513  
 A:Molecule type: DNA  
 A:Residues: 1-103 <RYA>  
 A:Cross-references: UNIPROT:P02458; UNIPROT:Q14042; UNIPROT:Q16672; UNIPROT:Q12985; UNIPROT:Q16084  
 R:Su, M.W.; Lee, B.; Ramirez, F.; Machado, M.; Horton, W.  
 Nucleic Acids Res. 17, 9473, 1989  
 A:Title: Nucleotide sequence of the full length cDNA encoding for human type II procollagen  
 A:Reference number: S06715; MUID:90067946; PMID:2587267  
 A:Accession: S06715  
 A:Molecule type: mRNA  
 A:Residues: 1-28, 'R', '99-1487' <SU2>  
 A:Cross-references: UNIPARC:UPI0000126D15; EMBL:X16468; NID:g29515; PIDN:CAA34488.1; PID  
 A>Note: alternative splice form 1  
 R:Vikula, M.; Mettaeranta, M.; Syvaenen, A.C.; Ala-Kokko, L.; Vuorio, E.; Peltonen, L.  
 Biochem. J. 285, 287-294, 1992  
 A:Title: Structural analysis of the regulatory elements of the type-II procollagen gene.  
 A:Reference number: S24270; MUID:92344585; PMID:1637314  
 A:Accession: S24270  
 A>Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-28 <VIK>  
 A:Cross-references: UNIPARC:UPI0000173B64; EMBL:X58709; GB:S40537; NID:g35659  
 A>Note: this translation is not annotated in Genbank entry HSPROCE01, release 111.0  
 R:Numaz, A.M.; Kohno, K.; Martin, G.R.; Yamada, Y.  
 Gene 44, 11-16, 1986  
 A:Title: Promoter region of the human pro-alpha-1-(II)-collagen gene.  
 A:Reference number: A24828; MUID:87031574; PMID:3021582  
 A:Accession: A24828  
 A:Molecule type: DNA  
 A:Residues: 1-8, 'T', '10-28' <NUN>  
 A:Cross-references: UNIPARC:UPI000016A71A; GB:M25698; NID:g180872; PIDN:AAAS2051.1; PID:  
 R:Baldwin, C.T.; Reginato, A.M.; Smith, C.; Jimenez, S.A.; Prockop, D.J.  
 Biochem. J. 262, 521-528, 1989  
 A:Title: Structure of cDNA clones coding for human type II procollagen. The alpha-1(II)  
 A:Reference number: S06496; MUID:90026318; PMID:2803268  
 A:Accession: S06496  
 A:Molecule type: mRNA  
 A:Residues: 7-28, 'R', '99-157', 'P', '159-440', 'G', '442-456', 'E', '458-640', 'A', '642-831', 'PA', '834', 'F'  
 A:Cross-references: UNIPARC:UPI0000173B65; EMBL:X16711; NID:g30040; PIDN:CAA34683.1; PID  
 A>Note: alternative splice form 1  
 R:Ryan, M.C.; Sandell, L.J.  
 J. Biol. Chem. 265, 10334-10339, 1990  
 A:Title: Differential expression of a cysteine-rich domain in the amino-terminal propept  
 A:Reference number: A35428; MUID:90285153; PMID:2355003  
 A:Accession: A35428  
 A>Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 27-81, 'L', '83-103' <RYA2>  
 A:Cross-references: UNIPARC:UPI0000173B66  
 A>Note: alternative splice form 2; splicing appears to be under developmental regulation  
 R:Su, M.W.; Benson-Chanda, V.; Vissing, H.; Ramirez, F.  
 Genomics 4, 438-441, 1989  
 A:Title: Organization of the exons coding for Pro alpha-1(II) collagen N-propeptide con  
 A:Reference number: A30147; MUID:89233138; PMID:2714801  
 A:Accession: A30147  
 A:Molecule type: DNA  
 A:Residues: 104-157, 'P', '159-236' <SUM>  
 A:Cross-references: UNIPARC:UPI0000173B67; GB:J03065; GB:M23660; GB:M25655; GB:M25656; G  
 R:Ala-Kokko, L.; Baldwin, C.T.; Moskowitz, R.W.; Prockop, D.J.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 6565-6568, 1990  
 A:Title: Single base mutation in the type II procollagen gene (COL2A1) as a cause of pri  
 A:Reference number: A94227; MUID:90370826; PMID:1975693  
 A:Accession: A33116

A:Molecule type: DNA  
 A:Residues: 171-172, 'C', '174-175' <ALA>  
 A:Cross-references: UNIPARC:UPI0000173B68  
 A>Note: mutant sequence from a family with family with primary generalized osteoarthritis  
 R:Diab, M.; Wu, J.J.; Eyre, D.R.  
 Biochem. J. 314, 327-332, 1996  
 A:Title: Collagen type IX from human cartilage: a structural profile of intermolecular  
 A:Reference number: S64673; MUID:96195147; PMID:8660302  
 A:Accession: S64674  
 A:Molecule type: protein  
 A:Residues: 188-189, 'X', '191-195', '1224-1230', 'X', '1232-1236' <DIA>  
 A:Cross-references: UNIPARC:UPI0000173B69; UNIPARC:UPI0000173B8A  
 R:Franc, S.; Marzin, E.; Boutillon, M.M.; Lafont, R.; Lechene de la Porte, P.; Herbage,  
 Eur. J. Biochem. 234, 125-131, 1995  
 A:Title: Immunohistochemical and biochemical analyses of 20000-25000-year-old fossil cal  
 A:Reference number: S63514; MUID:96096730; PMID:8529631  
 A:Accession: S63514  
 A:Molecule type: protein  
 A:Residues: 243-261; 575-590; 756-763, 'X', '765-779' <FRA>  
 A:Cross-references: UNIPARC:UPI0000173B6B; UNIPARC:UPI0000173B6C; UNIPARC:UPI0000173B6D  
 R:Tillier, G.E.; Weis, M.A.; Polunbo, P.A.; Gruber, H.E.; Rimoin, D.L.; Cohn, D.H.; Eyre,  
 Am. J. Hum. Genet. 56, 388-395, 1995  
 A:Title: An RNA-splicing mutation (G+5IVS20) in the type II collagen gene (COL2A1) in a  
 A:Reference number: I38867; MUID:95150028; PMID:7847372  
 A:Accession: I38867  
 A>Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 440, 'G', '442-456', 'E', '458-480', 'P', '482-509' <TIL1>  
 A:Cross-references: UNIPARC:UPI000006F3AF; EMBL:U15195; NID:g557053; PIDN:AAB60370.1; PI  
 R:Ramirez, F.  
 submitted to the EMBL Data Library, December 1988  
 A:Reference number: S04892  
 A:Accession: S04892  
 A:Molecule type: mRNA  
 A:Residues: 501-676, 'A', '678-783', 'A', '785-831', 'PA', '834', 'F', '836-1214' <RAM>  
 A:Cross-references: UNIPARC:UPI000016A700; EMBL:X13783; NID:g30037; PIDN:CAA32030.1; PI  
 R:Vikula, M.; Peltonen, L.  
 FEBS Lett. 250, 171-174, 1989  
 A:Title: Structural analyses of the polymorphic area in type II collagen gene.  
 A:Reference number: S05000; MUID:89325561; PMID:2753125  
 A:Accession: S05000  
 A:Molecule type: DNA  
 A:Residues: 630-640, 'A', '642-785' <VIK2>  
 A:Cross-references: UNIPARC:UPI0000173B6E; EMBL:X16158; NID:g29951; PIDN:CAA34278.1; PID:  
 A34282.1; PID:g1335022; PIDN:CAA34283.1; PID:g1335023; PIDN:CAA34284.1; PID:g1335024  
 R:Bogaert, R.; Tillier, G.E.; Weis, M.A.; Gruber, H.E.; Rimoin, D.L.; Cohn, D.H.; Eyre, D.  
 J. Biol. Chem. 267, 22522-22526, 1992  
 A:Title: An amino acid substitution (Gly853-->Glu) in the collagen alpha 1(II) chain pro  
 A:Reference number: A44309; MUID:93054548; PMID:1429602  
 A:Accession: A44309  
 A>Status: nucleic acid sequence not shown; not compared with conceptual translation  
 A:Molecule type: DNA; mRNA  
 A:Residues: 752-831, 'PA', '834', 'F', '836-1005', 'K', '1007-1036', 'Q', '1038-1052', 'E', '1054-1068', 'T',  
 A:Cross-references: UNIPARC:UPI0000173B6F; GB:L00977; NID:g180812; PIDN:AAB23914.1; PID:  
 A>Note: sequence extracted from NCBI backbone (NCBI:117273); parts of this sequence wer  
 A>Note: this translation is not annotated and this publication is not cited in GenBank e  
 A>Note: mutant sequence associated with perinatal lethal hypochondrogenesis  
 R:Tillier, G.E.; Rimoin, D.L.; Murray, L.W.; Cohn, D.H.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 3889-3893, 1990  
 A:Title: Tandem duplication within a type II collagen gene (COL2A1) exon in an individua  
 A:Reference number: S16502; MUID:90251662; PMID:2339128  
 A:Accession: S16502  
 A:Molecule type: DNA  
 A:Residues: 1164-1184, 'GPSKDGANGIPQPI', '1185-1199' <TIL2>  
 A:Cross-references: UNIPARC:UPI000011F7F2; EMBL:M37126; NID:g180808; PIDN:AAAS2037.1; PI  
 A>Note: mutant sequence from a patient with spondyloepiphyseal dysplasia  
 R:Cheah, K.S.E.; Stoker, N.G.; Griffin, J.R.; Grosveld, F.G.; Solomon, E.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 2555-2559, 1985  
 A:Title: Identification and characterization of the human type II collagen gene (COL2A1)  
 A:Reference number: A02858; MUID:85190534; PMID:3857598  
 A:Accession: A02858  
 A:Molecule type: DNA  
 A:Residues: 1032-1056, 'N', '1058-1068', 'T', '1070-1487' <CHE>





A:Introns: 54/3; 285/2

## Alignment Scores:

Pred. No.: 0.00974 Length: 301  
Score: 143.50 Matches: 68  
Percent Similarity: 28.4% Conservative: 10  
Best Local Similarity: 24.7% Mismatches: 92  
Query Match: 3.3% Indels: 105  
DB: 2 Gaps: 13

US-10-768-158-1 (1-2419) x T21314 (1-301)

QY 780 TCCACAGCCCACTCTTCCCGCCGCCACGGCAGGCGCTCAGCGT----- 736  
DB 80 ThrThrArgGlnAlaTyGlyGlyProGluValAsnProAlaProAsnLeuGlnCysGlu 99  
QY 735 -----TGCAAGCACTGGTCCACAGCTGGTGGCAGTGTCTCCGTCA----- 697  
DB 100 GlyCysCysLeuProGlyProGlyProAlaGlyAlaProGlyLysProGlyLysPro 119  
QY 696 GGGCTTCAGCTGGGCTTGTACAGGACACCCCGAGGAATCTGGCCAGCTGCTCCA--- 640  
DB 120 GlyArgProGlyAlaProGlyThrProGlyThrProGlyLysProProValAlaProCys 139  
QY 639 ---CCATCGTCACAGCTCCCATGCTCTTCATCTTGAGAAAACACGTTCCGAGT 583  
DB 140 GluProThrThrProProCysLysProCysProGlnGlyProProGlyProGly 159  
QY 582 CCATCGGTGCTCCAGAACTCTCGCAGCTGTCAAAACGAGGCGGTAGCCAGCTTAT 523  
DB 160 ProProGlyAlaProGlyAspProGlyGluAlaGlyThrProGlyArg----- 175  
QY 522 CATTCATAAACCTCCGGCAGAAATCTTGAAGGTGCTCGTAGCTCATGGTCCGAGAG 463  
DB 176 -----ProGlyThrAspAlaAla 181  
QY 462 AGCGGTGAGACTGATAAATAGACACACAGATGCTTGGGTGGAGCCATATAGATGA 403  
DB 182 ProGlySerProGlyProArgGlyProProGlyPro----- 193  
QY 402 CTTGGAGTCTCAATGTGGAGGTGAGAGGCGAGAAAGCGTAGGCGAGTGGCTCTTGA 343  
DB 194 -----AlaGlyGluAlaGly----- 198  
QY 342 TGAGCGGGGAGAGTCACTTCTTGATGATGCCAGGCCGGCTGTGGGTACTCCAGGA 283  
DB 199 -----AlaProGlyProAla---GlyGluProGly 207  
QY 282 CCGGAGGTGCTGCTGATGTTTCATCAACCGCATCTATCGGGGTACGGCCCTGCTCA 223  
DB 208 ThrProAlaIleSerGluProLeuThr-----ProGlyAla 219  
QY 222 CCAAGTAGACCA-----CCTCTGCAGCAAGCTGGTGCCGACT 184  
DB 220 ProGlyGluProGlyAspSerGlyProProGlyProProGlyPro----- 235  
QY 183 TGGGGTAGGTGACGATCCACACGCTGCTGGGCGGCACCGGAAGTTGGCGATCTCTCCA 124  
DB 236 -----GlyAlaProGlyAsn---AspGlyProPro 244  
QY 123 TCTTCCCGCGGAGAAAGCGGCGAGCCGACGCGCATFGAACTCGAAGTACTTGTCTCGA 64  
DB 245 GlyProProGlyProLysGlyAlaProGlyProAspGly----- 257  
QY 63 ACTCCCGCGGGTGTGGGG---TCTCGGCTCGCTCTCCGCCA 22  
DB 258 ---ProProGlyValAspGlyGlnSerGlyProProGlyProPro 271

RESULT 97

T20497

hypothetical protein F02D10.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T20497

R:Swinsburne, J.

submitted to the EMBL Data Library, November 1995

A:Reference number: Z19283

A:Accession: T20497

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-316 &lt;WIL&gt;

A:Cross-references: UNIPROT:Q19111; UNIPARC:UPI0000080924; EMBL:Z67990; PIDN:CAA91932.1.1

A:Experimental source: clone F02D10

C:Genetics:

A:Gene: CESP:F02D10.1

A:Map position: X

A:Introns: 56/3

## Alignment Scores:

Pred. No.: 0.00984 Length: 316  
Score: 143.50 Matches: 66  
Percent Similarity: 30.1% Conservative: 8  
Best Local Similarity: 26.8% Mismatches: 87  
Query Match: 3.3% Indels: 85  
DB: 2 Gaps: 13

US-10-768-158-1 (1-2419) x T20497 (1-316)

QY 735 TGCAGCACTGGTCCACAGCTGGTGGCAGTGTCTCCGTCA----- 697  
DB 117 CysAsnProGlyProProGlyProGlyArgProGlyLysProGlyThrProGlyLys 136  
QY 696 ---GGCTTCCA-----GCTGGGCTTGTCA--- 673  
DB 137 ProGlyAlaProGlyAsnProGlyAlaSerGlyLysGlyAlaAlaProCysGluPro 156  
QY 672 AGGACACCCCGAGGAATC-----TGGCCAGCTGTCCACCATCTGCCAGGTCCCGAT 619  
DB 157 SerThrProProCysGlnProCysProAlaGlyProProGlyProProGlyProAsp 176  
QY 618 GCATGCTTCTATCTTGAGAAAACACGTTTCGAGTCCATGCGTGTCTCCAGAACCTCT 559  
DB 177 GlyThr-----ProGlyGluProGlyProGlyProAlaGlySerProAlaGlyPro 194  
QY 558 GCACGTGCTCAACCCAGGAGCGGTAGCCAGCTTATCATTAACCTCCGGCAGAAAT 499  
DB 195 Ser----- 195  
QY 498 CTTGAAAGGTGCTCGGTAGTCTCATGTCCGAGAGCGGTGGAACATATAAAGACA 439  
DB 196 -----GlyProGlyProAlaGly 201  
QY 438 CCACAGATCTTGGGTTGCGAGCCATATAGATGACCTTGGAGTCTCCATTTGGAGGT 379  
DB 202 ProProGlyProAlaGly-----Pro-----ProGlyAsnAspGlyGlnProGly 216  
QY 378 CAGAGGCGAGAAAGCGGTAGGCGAGTGGCTTGTATGAGGCGGGGAGAGGTGAGTCTCT 319  
DB 217 GlnProGlyGlyProGlyGlnAspGlyAlaSerSerAlaGlyGlyGluAlaGly----- 234  
QY 318 TGATGATGTCAGAGCGCGCTGTGGTACTCCAGGACCG---GGAGCTGCTCGTCGATGT 262  
DB 235 -----ProGlyProAla---GlyProProGlyProAlaGlyProAlaGlyProAsp 250  
QY 261 TCATCAACCCGATCTCATCGGGGTGAGCGCTTGGCTCCCAAGTAGACCTCTCTGCA 202  
DB 251 GlyGlnSerGlySerGlySerAlaGlyGlyProGlyProLysGlyProProGlyProAla 270  
QY 201 GCAAGCTGTGCGGACTTGGGGTAGGTGACGATCCACACGTCGCTGGCGCGCACCGGA 142  
DB 271 GlyGln-----ProGly----- 274  
QY 141 AGTTGGCGATCTCTCTCCATCTTCCGCGCGCAGAGGGCGGCGGACCGCATGGAAC 82  
DB 275 Ser---AspGlyAsnProGlyThrAlaGlyProProGlyAsnProGlyGlyGluGlyGlu 293



Score:	143.50	Matches:	82
Percent Similarity:	31.6%	Mismatches:	11
Best Local Similarity:	27.9%	Indels:	99
Query Match:	3.3%	Gaps:	18
DB:	1		
US-10-768-158-1 (1-2419) x CGRTIS (1-671)			
QY	762	CCCCGCCACGGGAGGCGCTCAGCGTTGCAGCACACTGGTCCACCAGCT-	715
DB	289	ProGlyPro--AlaGlyValGln-----GlyProGlyProAlaGlyGlu	303
QY	714	-----GGTGGCAGTGCTCCGTCAGGGCTTCAGCTGGGCGCTTGT	676
DB	304	GluGlyLysArgGlyAlaArgGlyGluProGlyProSerGlyLeuProGlyProGly	323
QY	675	CACAGGACACCCCCAGCAATC-----	655
DB	324	GluArgGlyProGlySerArgGlyPheProGlyAlaAspGlyValAlaGlyProLys	343
QY	654	TGGCCAGCTGCTCCACCATGCTCACCAGGTCCCGATGCATGCTCTTCACTTGA	595
DB	344	GlyProAlaGlyGluArgGlySerProGlyPro-----	354
QY	594	GCAGTTCGAGTCCATCGGTGCTCCAGAACTCCTGCAGTGTCTAAACACGAGGCGGT	535
DB	355	-----AlaGlyProLysGlySerProGlyGluAlaGlyArgProGlyGluAlaGly	371
QY	534	AGCCAGCTTATCATTTATAAACCTCCGGCAGAAATCTTGAAAGGTGCCTCGGTAGTCA	475
DB	372	-----LeuProGlyAlaLysGlyLeuThr	379
QY	474	TGTTCCGACAGAGCGGTGGAAGTATAAAGACA---CCACCAGATCCTTGGGGTTC	418
DB	380	GlySerProGlySerProGlyProAspGlyLysThrGlyProProGlyProAlaGly	398
QY	417	GAGCCATATAGATGACCTTGGAGTCTCCATTGTGGAGGTGAGAGGGCAGAAAGC-	364
DB	399	*****GlyArg---ProGlyProAlaGlyProProGlyAlaArgGlyGlnAlaGlyVal	417
QY	363	-----GGTAGGCGAGGTGGCTTTGATGAGGCGGGGAGAGGTCA	325
DB	418	MetGlyPheProGlyProLysGlyThrAlaGlyGluProGlyLysAlaGlyGluArgGly	437
QY	324	GTTCTTTGATGATGTTCCAGGCCCC-	301
DB	438	Val-----ProGlyProProGlyAlaValGlyProAlaGlyLysAspGlyGlu	453
QY	300	-----CCTGTGGTACTCCAGACCGGGAGTCTCTCGTGATGTTTCATCAGCCGATCT	247
DB	454	AlaGlyAlaGlnGlyAlaProGlyProAlaGlyPro-----AlaGlyGluArgGly	470
QY	246	CATCGGGGTGAG-----CGCCCTGGCTCACCAAGTAGACACCTCTCGCAGCAAGCTGG	193
DB	471	GluGlnGlyProAlaGlySerProGlyPheGlnGlyLeuProGlyProAlaGly-----	488
QY	192	TGCCGGAAGTTGGGTAGGTGACGATCCACAGTCTGGTGGGCGCCACCGGGAAGTTGGCGA	133
DB	489	-----ProProGlyGluAlaGlyLysProGly*****GlyValProGlyAspLeuGly	506
QY	132	TCTCTCT---CCATCT---TCCCGGGGAGAAAGGGCGGCGCAGCCCATGGAACTCGA	79
DB	507	AlaProGlyProSerGlyAlaArgGlyGluArgGlyPheProGlyGluArgGlyValGln	526
QY	78	AGTACTTGCTCTCGAACTCCCGGGGTGCTGGGGGTCTCGG	37
DB	527	Gly-----ProProGlyProAlaGlyProArg	535

N;Alternate names: ORF 1 protein  
Nypotnetical 119.5K protein (uvrA region) - Micrococcus luteus  
C;Species: Micrococcus luteus.  
Micrococcus lysodeikticus

C>Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 04-Feb-2000  
C/Accession: JQ0405  
R/Shiota, S.; Nakayama, H.  
Mol. Gen. Genet. 217, 332-340, 1989  
A/Title: Micrococcus luteus homolog of the Escherichia coli uvra gene: identification of  
A/Reference number: S04781; MUID:89364717; PMID:2549377

A/Accession: JQ0405  
A/Molecule type: DNA  
A/Residues: 1-1106 <SHI>  
A/Cross-references: UNIPARC:UPI0000177394; EMBL:X15867  
A/Note: this reading frame extends between two stop codons and does not begin with a sta  
A/Note: the gene encoding this protein overlaps uvra gene  
C/Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

Alignment Scores:  
Pred. No.: 0.0129 Length: 1106  
Score: 143.50 Matches: 102  
Percent Similarity: 34.2% Conservative: 17  
Best Local Similarity: 29.3% Mismatches: 91  
Query Match: 3.3% Indels: 139  
DB: 2 Gaps: 22

US-10-768-158-1 (1-2419) x JQ0405 (1-1106)

```
QY 5 CGCGCAGCGCGCGCATGCGGAGAGCGAGCGAGCCCGCAGACCCCGCGGGAGTT 64
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
317 ArgValArgGlyArgGlyAlaGlyArgArgGluGlyGlnGluHis---ArgGlyVal 335
QY 65 CGAGAGCAAGTACTTCGA-----GTTCCATGCGGTGCGGCTGCCGCCCTTCTG 112
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
336 ArgArgAlaGlyArgGlyGlnProValProLeuValLeuArgGluAlaLeuLeu 355
QY 113 CGCGGGAGAGTAGGAGAGATCGCAACTTCCCGTGGCGCC-----CAGCGA 160
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
356 ProGlnArgAlaArgAlaAspArgGlyArgAspArgAlaAlaLeuValLeuGlnGln 375
QY 161 CQTGTGGATCGNCACCTACCCCAAGTC---CGGACCCAGCTTGTGCGAGGAGTGTCTA 217
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
376 ProVal---ArgArgValProArgValHisArgHisArgLeuProProAlaGlyGly--- 393
QY 218 CTTGTGTAGCGGCGGTGACCCCGATGAGATCGGCTTGATGAACATCGACGACGCT 277
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
394 -----ProGly-----ProArg-----ArgArgGlnArgAlaVal 404
QY 278 CCGGTT-----ProArg-----ArgArgGlnArgAlaVal 404
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
405 ProAlaArgGlyArgArgAlaValValAlaArgGlnValHisLeuGlyLeuAla 424
QY 284 -----CCTGGAGTACCCACAGCGCGG 304
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
425 AlaArgAlaArgAlaGlyGlnGlyAspGlyLeuProGlyHisProValGluGly 444
QY 305 CCT-----GGACATCATCAAGGAAGTGC 328
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
445 ProAspGlyGlyAlaArgArgProAlaArgGlnGlyLeuGlnGlyGlyAsp 464
QY 329 CTCCTCCCGCCTCATCAGAGCCACCTGCCCTACCGCTTCTGCCCTCTGACCTCCAA 388
Db ::|||::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
465 Val-ProGlnProValArgProAlaPro-----LeuHis 477
QY 389 TGGAGACTCCAAGTCTATATGGCTGCGCAACCCCAAGGATCTGGTGTCTTATTA 448
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
477 eGly-----LeuArgGlyArgHisPro----- 484
QY 449 TCAGTTCCACCGCTCTTCGGACCATGAGCTACCGAGCACCTTTCAAGAATTCGCG 508
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
485 -----LeuArgAspAlaGlnAlaArgGly----- 492
QY 509 GAGGTTTATGATGATACTGGGCTA-----CGGCTCTGTTTGGACACGTGCAGG 561
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
493 -----AspArgValGlyArgArgProArgAlaLeuArgValValHisAlaG 508
QY 562 A-----GTTCTGGGAGCA-----CCGCATGGACTCGAACGTGCTTT 597
```

```
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
508 YAspProValProGlyValProArgGlyProProGlnProHisGlyProGluArgAlaAr 528
QY TTCTCAA---GTATGAAGA-----CATGCATCGGACCTGGT----- 631
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
528 gGlyArgProValHisArgGlyArgHisProProAlaHisAlaArgGlyHisGlyVal 548
QY 632 -----GACGATGTTGGAGCAGCTGCCAGATTCTCTGGGGTGTCTCTGTG 675
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
548 uLeuGlyAlaAlaAlaAspGlyProGlyAlaAlaAspArgGlyProGlyAlaGlnGlyAs 568
QY 676 ACAAGGCCACGCTGGAAGCCCTGACGGAGCACTGCCA-----CCAGC 717
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
568 pProGlyProAlaGlyValProAlaGlyArgArgProArgValProGlnProArgAlaAl 588
QY 718 TGGTGGACCATGCTGCAACGCTGAGCCCTGCGGCGCGGGAAGAGATTGGGCTGT 777
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
588 aGlyArgHisProLeuArgArgArgGlyProAlaHisProPro----- 602
QY 778 GGAAGGACATCTTCACCGTCTC 799
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
603 ----GlyHisThrAspArgLeu 608
```

Search completed: May 17, 2006, 10:00:17  
Job time : 252.5 secs



```
Sequence 31, Appl
Sequence 32, Appl
Sequence 87, Appl
Sequence 13, Appl
Sequence 518, App
Sequence 97, Appl
Sequence 40, Appl
Sequence 393, App
Sequence 97, Appl
Sequence 2, Appl
Sequence 2466, Ap
Sequence 7, Appl
Sequence 17, Appl
Sequence 54, Appl
Sequence 2959, Ap
Sequence 466, App
Sequence 21, Appl
Sequence 224, App

83      83      1.9      558      7      US-11-311-754-31
84      83      1.9      811      6      US-10-505-928-32
85      83      1.9      811      6      US-10-505-928-87
86      83      1.9      951      6      US-10-199-229-13
87      83      1.9      1137     6      US-10-196-749-518
88      83      1.9      1894     6      US-10-196-749-97
89      82.5     1.9      665      6      US-10-370-959-40
90      82.5     1.9      667      6      US-10-505-928-393
91      82.5     1.9      1894     6      US-10-196-749-97
92      82.5     1.9      4373     7      US-11-118-524-2
93      82      1.9      362      6      US-10-511-937-2466
94      82      1.9      872      6      US-10-519-335-7
95      82      1.9      872      7      US-11-302-678-17
96      81.5     1.9      384      6      US-10-505-928-54
97      81.5     1.9      620      6      US-10-511-937-2959
98      81.5     1.9      813      6      US-10-196-749-466
99      81.5     1.9      1763     6      US-10-504-120-21
100     81      1.8      341      6      US-10-196-749-224

ALIGNMENTS

RESULT 1
US-10-985-570-2
; Sequence 2, Application US/10985570
; Publication No. US20060100138A1
; GENERAL INFORMATION:
; APPLICANT: Olsen, David R.
; APPLICANT: Polarek, James W.
; APPLICANT: Yang, Chunlin
; TITLE OF INVENTION: IMPLANTABLE COLLAGEN COMPOSITIONS
; FILE REFERENCE: FP0305 US
; CURRENT APPLICATION NUMBER: US/10/985.570
; CURRENT FILING DATE: 2004-11-10
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-985-570-2

Alignment Scores:
Pred. No.: 9,13e-06 Length: 1464
Score: 155.50 Matches: 86
Percent Similarity: 29.8% Conservative: 11
Best Local Similarity: 26.5% Mismatches: 107
Query Match: 3.5% Indels: 121
DB: 6 Gaps: 20

US-10-768-158-1 (1-2419) x US-10-985-570-2 (1-1464)

QY 783 CTTTCCACAGCCCACTCTCCCGGCCCA-----CGGCAGGCGCTCAGGTTGCAGC 730
Db 870 ProProGlyAlaThrGlyPheProGlyAlaAlaGlyArgValGlyProProGlyProSer 889
QY 729 -----ACTGTCACACAGCT-----GGTGGCAGTGCTCCG 700
Db 890 GlyAenAlaGlyProProGlyProProGlyProAlaGlyLysGluGlyGlyLysGlyPro 909
QY 699 TCAGGCTTCCAGCTGGGCGCTGTGCACAGACA-----CCCCCA 661
Db 910 ArgGly-----GluThrGlyProAlaGlyArgProGlyGluValGlyProProGlyPro 928
QY 660 GGAATCTGGCAGCTGCTCCACCATCGTCACAGGTCCTGCCAGACTCTCGACGTGCTCAACACAGG 541
Db 929 Gly-----ProAlaGlyGluLysGlySerProGlyAlaAsp----- 940
QY 600 GAAAGACAGGTTCCAGTCCATGCGGTGCTGCCAGACTCTCGACGTGCTCAACACAGG 541
Db 941 -----GlyProAlaGlyAlaProGlyThrProGlyProGluGlyLysAla 955
QY 540 AGCCGTAGCCAGCTTATCATTTATAAACCTCCGCGAGAATTTCTTGAAGAGTGCTC--- 484
```

```
DB: 6 Gaps: 16
US-10-768-158-1 (1-2419) x US-10-985-570-1 (1-1466)
QY 765 TTCCCGCCGACCGGACGGCCTCAGCGTTGCAGCACTGCTCCACAGCT-----GGT 712
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 172 TyrProGlyPro---AlaGlyPro-----ProGlyProProGlyProProGly 186
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 711 GGCACTGCTCGCTCAGGCTTCCAGCTGGGCTTTGCAGGACACACCCCGGAGG----- 658
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 187 ThrSerGlyHisProGlySerProGlySerProGlyTyrGlnGlyProProGlyGluPro 206
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 657 -----ATCTGGCCAGCTGCTCCACCATCGTCACAGGTCCCGATGCGATGCTTCATACT 604
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 207 GlyGlnAlaGlyProSerGlyProProGlyProProGly----- 219
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 603 TGAGAAAAGCACGTTGAGTCCATGCGGTGCTCCAGAACTCTCGCAGCTGCTCAAAACC 544
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 220 -----AlaIleGlyProSerGlyProAlaGlyLysAspGly 231
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 543 AGGAGCGGTAGCCAGCTTATCATTTATAAACCTCCGGCAGAAATTTTGAAGGTGCTC 484
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 232 GluSerGlyArgPro-----GlyArgProGlyGluArgGlyLeu 244
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 483 GGTAGCTCATGTCGCGAGAGCGGTGGAACCTGATATAAGACACACCATGATCTTGG 424
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 245 -----ProGlyProPro 248
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 423 GGTTCGAG---CCATATAGATGACCT-----TGGAGTCTCCATTGTGGA--- 382
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 249 GlyIleGlyProAlaGlyIleProGlyPheProGlyMetLysGlyHisArgGlyPhe 268
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 381 ---GGTCAGAGGCGAGAAAGCGGTAGGCGAGGTGGCTTTGATGAGGCGGAGAGGTCA 325
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 269 AspGlyArgGlnGlyGluLeuGlyGluThrGlyAlaProGlyLeuLysGlyGluAsnGly 288
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 324 GTTCCT-----TGATGATGTCAGGCCCG----- 301
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 289 LeuProGlyGluAsnGlyAlaProGlyProMetGlyProArgGlyAlaProGlyGluArg 308
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 300 ---CCTGTGGTACTCCAGGACCGGAGCTGCTCGTGTGATTTTCATCAAGCGCATCTCAT 244
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 309 GlyArgProGlyLeuProGlyAlaAlaGlyAlaArgGlyAsnAspGlyAlaArgLysSer 328
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 243 CGGGGTGAGCCCTCGCTCCAAAGTAGACCACTCTCTGCAAGCTGCTGCGGACT 184
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 329 AspGlyGln---ProGlyProProGly---ProProGlyThrAlaGlyPheProGlySer 346
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 183 TGGGGTAGTGACGATCCACACGTCGCTGGGCCGACCGGAAGTTGGCGATCTCTCCA 124
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 347 ProGlyAlaLysGlyGluValGlyProAlaGlySerProGlySer---AsnGlyAlaPro 365
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 123 TCTTCCCGCGCAAGGGCGGCGCCGACCGCATCGAACTCGAAGTACTTGTCTCTCGA 64
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 366 GlyGlnArgGlyGluProGlyProGlnGlyHisAlaGlyAlaGlnGlyProProGly--- 384
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 63 ACTCCCGCGGGGTGCTGGGGGTCT 40
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 385 ---ProProGlyIleAsnGlySer 391
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
US-10-985-570-3
; Sequence 3, Application US/10985570
; Publication No. US20060100138A1
; GENERAL INFORMATION:
; APPLICANT: Olsen, David R.
; APPLICANT: Polarek, James W.
; APPLICANT: Yang, Chunlin
; TITLE OF INVENTION: IMPLANTABLE COLLAGEN COMPOSITIONS
; FILE REFERENCE: FP0305 US
; CURRENT APPLICATION NUMBER: US/10/985,570
; CURRENT FILING DATE: 2004-11-10
; NUMBER OF SEQ ID NOS: 3
```

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-985-570-3
Alignment Scores:
Pred. No.: 0.000158 Length: 1366
Score: 140.50 Matches: 84
Percent Similarity: 28.5% Conservative: 11
Best Local Similarity: 25.2% Mismatches: 97
Query Match: 3.2% Indels: 141
DB: 6 Gaps: 19
US-10-768-158-1 (1-2419) x US-10-985-570-3 (1-1366)
QY 840 CCATCTCTCTTTATACACCAAGTCAAACTCTCATTCATGAGACGGTGAAGATGTCCT 781
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 204 ProGlyGluAsnGlyThrProGlyGlnThrGly----- 214
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 780 TCCACAGCCCAACTCTTCCCGGCCACGG----- 751
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 215 -----AlaArgGlyLeuProGlyGluArgGlyValGlyAlaProGlyProAlaGly 232
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 750 -----GCAGGCTCTACGCTTGCAGC 730
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 233 AlaArgGlySerAspGlySerValGlyProValGlyProAlaGlyProAsnGly---Ser 251
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 729 ACTGCTCCACA-----GCTGGT 712
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 252 AlaGlyProProGlyPheProGlyAlaProGlyProLysGlyGluIleGlyAlaValGly 271
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 711 GGAGTCTCGCTCAGGCTTCCAGCTGGGCTTGTACAGGACACCC---CCAGGATC 655
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 272 AsnAlaGlyProThrGly---ProAlaGlyProArgGlyGluValGlyLeuProGlyLeu 290
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 654 ---TGGCCAGCTGCTCCACCATCGTCAACAGTCCCATGTCATCTTCTACTTGAAGAA 598
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 291 SerGlyProValGlyProProGlyAsnProGlyAlaAsnGly---LeuThrGlyAlaLys 309
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 597 AAGCAGCTTCGAGTCCA-----TGGCGTCTCCAGAACTCCT----- 559
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 310 GlyAlaAlaGlyLeuProGlyValAlaGlyAlaProGlyLeuProGlyProArgGlyIle 329
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 558 -----GCAGTGTCTCAACACGAGCGGTAGC 532
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 330 ProGlyProProGlyAlaAlaGlyThrThrGlyAlaArgGlyLeuValGlyGluProGly 349
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 531 CCAGCTTATCATTAACCTCCGCGAAGATTTCTTGAAGGTGCTCGGTAGTCTATGG 472
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 350 ProAlaGlySerLysGlyGluSerGlyAsn----- 359
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 471 TCCGACAGAGCGGTGGAAGTATATAAGACACCAACAGATCTCTGGGTTCGAGCCA 412
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 360 LysGlyGluProGlySerAlaGlyProGlnGlyProProGlyProSerGlyGluGlu--- 378
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 411 TATAGATGACTTGGAGTCTCCATTGTGGAGGTGAGAGGCGAGAAAGCGGTAGGCGAGGT 352
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 379 -----GlyLysArgGlyProAsnGlyGluAlaGly 388
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 351 GGCTTTGATGAGCGGGGAGAGGTGAGTTCTTGATGATGTCCAGGCCCGCGGTGGGT 292
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 389 -----SerAlaGly 391
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 291 ACTCCAGGACGGGAGCTGCTCGTCGATGTTTCATCAAGCCGATCTCATCGGGTCAAGCG 232
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 392 ProProGlyProProGlyLeuArg----- 399
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 231 CTGGGCTCACCA---AGTAGACCACTCTCTGACAGCAAGCTGGTCCGCGACTTGGGTAGG 175
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 400 ---GlySerProGlySerArgGlyLeuProGlyAlaAsp-----GlyArg 413
Db :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
QY 174 TGACGATCCACACGTCGCTGGCGCGCACCGGGAAGTTGGCGATCTCTCCATCTTCCGC 115
Db 414 -----AlaGlyValMetGlyProGlySerArg 423
QY 114 GGCAGAAAGGGCGGAGCGCCGCGCATGGAACTCGAAGTACTTCTCTCGAACT---CCC 58
Db 424 GlyAlaSerGlyProAlaGlyValArgGlyProAlaGlyAspAlaGlyArgProGlyGlu 443
QY 57 CCGGGGTGCTGGGGTCTCGG-----CCTCGCTCTCGG 25
Db 444 ProGlyLeuMetGlyProArgGlyLeuProGlySerPro 456
RESULT 4
US-10-985-570-2
; Sequence 2, Application US/10985570
; Publication No. US20060100138A1
; GENERAL INFORMATION:
; APPLICANT: Olsen, David R.
; APPLICANT: Polarek, James W.
; APPLICANT: Yang, Chunlin
; TITLE OF INVENTION: IMPLANTABLE COLLAGEN COMPOSITIONS
; FILE REFERENCE: FP0305 US
; CURRENT APPLICATION NUMBER: US/10/985.570
; CURRENT FILING DATE: 2004-11-10
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-985-570-2
Alignment Scores:
Pred. No.: 0.000622 Length: 1464
Score: 133.50 Matches: 72
Percent Similarity: 32.2% Conservative: 12
Best Local Similarity: 27.6% Mismatches: 73
Query Match: 3.0% Indels: 104
DB: 6 Gaps: 15
US-10-768-158-1 (1-2419) x US-10-985-570-2 (1-1464)
QY 19 GCATGGCGGAGCGAGCGAGCCAGCCAGCCCGGGGGAGTTCGAGAGCAAGTACT 78
Db 508 AlaGlyGluArgGlySerProGlyProAlaGlyProLysGlySer----- 522
QY 79 TCGAGTTCATGGCTGGCGTGGCGCCCTTCTGCGCGGGAAGATGGAGATCGCCA 138
Db 523 -----ProGlyGluAlaGlyArg-ProGlyGluAlaG1 533
QY 139 ACTTCCCGGTGGGCCAGCAGCGTGTGGATCGTCACTACCCCAAGTCCCGCACCAGCT 198
Db 533 yLeuProGlyAlaLys----- 538
QY 199 TGCTGCAGGAGTGTCTTACTTGTGAGCGAGGCGCTGACCCCGATGAGATCGGCTGA 258
Db 539 -----GlyLeuThrGlySerProGly----- 545
QY 259 TGAACATCGACGAGCAGCTCCCGGTCTGGAGTACCCACAGCCGGCGCTGACATCA 318
Db 546 -----SerProGlyProAspGlyLysThrGlyProProGly---ProAl 559
QY 319 AGGAAGTACCTCTCCCGCTCATCAGAGCCACCTCCCTACCGCTTCTGCGCCTCG 378
Db 559 aGlyGlnAspGlyArgPro-----GlyProProGly----- 569
QY 379 ACCTCCCAATGAGACTCCCAAGTCACTATATGGC---TCGCAACCCCAAGGATCTGG 435
Db 570 -ProProGlyAlaArgGlyGlnAlaGlyValMetGlyPheProGlyProLysGlyAlaAl 589
QY 436 TGGTGTCTTATATCAGTTCACCGCTCTCTGCGGAGCACCATGAGCTACCGAGCCTTTC 495
Db 589 aGly-----GluProGlyLysAlaGlyGluArgGlyValProGlyProProG1 605
```

```
QY 496 AAGAAATTCCTCCGAGGTTTATGAATGATAAGCTGGGCTACGGCTCCTGTTGAGCACG 555
Db 605 yAlaValGlyPro-----AlaGlyLysAsp-----GlyG1 615
QY 556 TGCAAGAGTTCCTGGAGCACCGCATGACTCGAAGCTGCTTTTCTCAAGTATGAGACA 615
Db 615 uAlaGlyAlaGlnGlyProPro----- 622
QY 616 TGCATCGGACCTGTGTCACGATGTGTGAGCAGCTGGCCAGATTCCTGGGGGTGTCCTGTG 675
Db 623 -----GlyProAlaGlyProAlaGlyGluArgGlyGlu----- 633
QY 676 ACAGGCCAGCTCGAAGCCCT-----GACGGAGCACTGCCACCAGCTGCTGGACCACT 729
Db 634 -GlnGlyProAlaGlySerProGlyPheGlnGlyLeuProGlyProAlaGlyProPro-- 652
QY 730 GCTGCAACGCTGAGGCCCTCCCGTGGCGCGGGAAGAGTGTGGCTCTGGAAGACATCT 789
Db 653 -----GlyGluAlaGlyLysProGlyGluGln--GlyValProGlyAspLe 667
QY 790 T 790
Db 667 u 667
RESULT 5
US-10-505-928-150
; Sequence 150, Application US/10505928
; Publication No. US20060089532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505.928
; CURRENT FILING DATE: 2004-08-27
; PRIOR FILING DATE: 2004-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 150
; LENGTH: 5738
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-150
Alignment Scores:
Pred. No.: 0.0171 Length: 5738
Score: 119.00 Matches: 88
Percent Similarity: 26.2% Conservative: 14
Best Local Similarity: 22.6% Mismatches: 115
Query Match: 2.7% Indels: 172
DB: 6 Gaps: 21
US-10-768-158-1 (1-2419) x US-10-505-928-150 (1-5738)
QY 786 TGCTCTCCACAGCCCAACTCTTCCCGGCCAGCGGCGCTCAGCGTTGCAGCACT 727
Db 1916 CysValSerProGluGlnLeuCysAspGlyValArgGlnCysProAspGlySerAspGlu 1935
QY 726 GGTCACCAAGTGTGGCAGTGTCCCGTCAGGGCTTCCAGCTGGGCGCTTGTACAGGACA 667
Db 1936 GlyProAspAlaCysValGluAla-----ProAlaProProAlaMetArgGly 1951
QY 666 CCCCAGGAATCTGGCAGCTGCTCCACCATCGTCACCAAGTCCCGATGCTGCTTCAT 607
Db 1952 ProProGly---GlnAlaGlyGlyProThrSerArgAlaPro----- 1965
QY 606 ACTTGAGAAAAGCACGTTTCGAGTCCATGCGGTCTCCAGAACTCTCCAGAACTCTTCAC 547
Db 1966 -----SerProProSerProProGluAlaGln 1974
QY 546 ACCAGGAGCCGTAGCCCGAGCTTATCATTAACCTCCCGCAGAAATT----- 499
Db -----
```



```
Db 1975 GlyGluGlyArgLysGlyGlnGluArgSerArgThrHisLeuThrValProAlaGlySer 1994
Qy 498 -----CTTGAAGGTGCTCGTAGCTCATGTCGCGAGAGCGGTGGAACATGATAAT 445
Db 1995 ThrGlnLeuProLeuCysProGlyLeuPheProCysGlyValAlaProGlyLeu---Cys 2013
Qy 444 AAGACACCA-----CCAGAT---CCTTGGGGTTGCGAGCCATAT 409
Db 2014 LeuThrProGluGlnLeuCysAspGlyIleProAspCysProGlnGlyGluAsp----- 2031
Qy 408 AGATGACCTTGGAGTCTCATTTGGAGGTGAGAGGCGAGAAAGCGGT-----AGGSCA 355
Db 2032 -----GluLeuAspCysGlyLeuProAlaLeuGlyGlyProAsnArgThr 2047
Qy 354 GGTGGCTTTGATGAGCGGGGAGAGGTTCAGTTCTTGATGATGTCCA----- 307
Db 2048 Gly-----LeuPro-----CysProGluTyrThrCys 2056
Qy 306 -----GGCCCGGCTGTG-----GGTACTCCAGGACCGGAGCT 274
Db 2057 ProAsnGlyThrCysIleGlyPheGlnLeuValCysAspGlyGlnProAspCysGlyArg 2076
Qy 273 GCTCGTCGATTTTCATCAAGCCGATCTCATCGGGTCAGCCCTGGCTCACCAAGTAGA 214
Db 2077 ProGlyGlnValGlyProSerProGluGlnGlyCysGlyAlaTrpGlyProTrpSer 2096
Qy 213 CCA---CCTCTCGCAGCAAGCTGTGTCGGGACTTGGGGTAGGTACGATCCACACAGT--- 160
Db 2097 ProTrpGlyProCysSerArgThrCysGlyProTrpGlyGlnGlyArgSerArgArgCys 2116
Qy 160 ----- 160
Db 2117 SerProLeuGlyLeuLeuValLeuGlnAsnCysProGlyProGluHisGlnSerGlnAla 2136
Qy 159 -----CGCTGGCGCCGACCGGGAGTGGCGA--- 133
Db 2137 CysPheThrAlaAlaCysProValAspGlyGluTrpSerThrTrpSerProTrpSerVal 2156
Qy 132 -----TCTCTCTCCATCTTCC 118
Db 2157 CysSerGluProCysArgGlyThrMetThrArgGlnArgGlnCysHisSerProGlnAsn 2176
Qy 117 CGCGCGAGAGGGCGGAGCC-----GCAGCCCATGGAACCTGGAAGTACT----- 73
Db 2177 GlyGlyArgThrCysAlaAlaLeuProGlyGlyLeuHisSerThrArgGlnThrLysPro 2196
Qy 72 -----TGC 70
Db 2197 CysProGlnAspGlyCysProAsnAlaThrCysSerGlyGluLeuMetPheGlnProCys 2216
Qy 69 TCTCGAATCTCC----- 58
Db 2217 AlaProCysProLeuThrCysAspAspIleSerGlyGlnValThrCysProProAspTrp 2236
Qy 57 -----CCGGGTGCTGG-----GGGTCTCG----- 37
Db 2237 ProCysGlySerProGlyCysTrpCysProGlyGlnValLeuGlySerGluGlyTrp 2256
Qy 36 -----CCTCGCTCTCCGCCATGC 19
Db 2257 CysValTrpProArgGlnCysProCys 2265
RESULT 6
US-10-985-570-3
; Sequence 3, Application US/10985570
; Publication No. US20060100138A1
; GENERAL INFORMATION:
; APPLICANT: Olsen, David R.
; APPLICANT: Polarek, James W.
; APPLICANT: Yang, Chunlin
; TITLE OF INVENTION: IMPLANTABLE COLLAGEN COMPOSITIONS
; FILE REFERENCE: FP0305 US
; CURRENT APPLICATION NUMBER: US/10/985.570
```

```
; CURRENT FILING DATE: 2004-11-10
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-985-570-3
Alignment Scores:
Pred. No.: 0.0174 Length: 1366
Score: 116.00 Matches: 80
Percent Similarity: 31.2% Conservative: 13
Best Local Similarity: 26.8% Mismatches: 78
Query Match: 2.6% Indels: 127
DB: 6 Gaps: 20
US-10-768-158-1 (1-2419) x US-10-985-570-3 (1-1366)
Qy 43 CCCCCAGCACCCGGGGAGTTTCGAGAGCAAGT-----ACTTCGAGTTTCATGGCGTGC 96
Db 293 ProValGlyProProGlyAsnProGlyAlaAsnGlyLeuThrGlyAlaLysGlyAlaAla 312
Qy 97 GGCTGCCGCCCTTCTCGCGGGAGAGATGGAGGAGATGCCCAACTTCCCGTGGCGGCCA 156
Db 313 GlyLeu---ProGlyValAlaGly-----AlaPro 321
Qy 157 GCGAGCGTGTGATCGTCACTTACCCTACCCCAAGTCCGGCACCCAGCTTCTGTCAGGA--- 208
Db 322 GlyLeuProGlyProArgGlyIleProGlyPro---ProGly-AlaAlaGlyThrGlu 340
Qy 209 ----GGTGGTCTACTTGGTGAGCCAGCGCGCTGACCCCGATGAGATCGGTTGATGAACA 264
Db 340 YAlaArgGlyLeuValGlyGluProGly-----ProAlaGlySerLysGlyGluSerG1 358
Qy 265 TCGACGAGCAGCTCCCGTCTCTGGAGTACCCACAGCGCGGCTGACATCATCAAGGAC 324
Db 358 YAsnLysGlyGluProGlySerAlaGlyProGlnGlyProProGlyProSerGlyGluG1 378
Qy 325 TGACCTCTCCCGCTCATCAAGAG-----CCACCTGCCCTACCGCTTTC 369
Db 378 uGlyLysArgGlyProAsnGlyGluAlaGlySerAlaGlyProGly----- 394
Qy 370 TGCCCTCTGACCTCCACAATGGAGACTCCAAGGTCTATATGGCTCGCAACCCCAAGG 429
Db 395 -----ProGlyLeuArg-----G1 400
Qy 430 ATCTGGTGTGCTTTATTATTCAGTTCCACCGCTCTCTGCGGACCATGAGTACCGAGGCA 489
Db 400 YSerProGly----- 403
Qy 490 CTTTTCAGGAATTCCTCGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCCTGTTTG 549
Db 404 ----SerArgGlyLeuProGlyAlaAsp----- 411
Qy 550 AGCAGCTGCAGGAGTCTTGGGAGCACCGCATGGACTCGAAGCTGCTTTTCTCAAGTATG 609
Db 412 -GlyArgAlaGlyValMetGlyProProGlySer----- 422
Qy 610 AAGCATGATCGGGACCT-----GGTGACGATGCT----- 640
Db 423 -ArgGlyAlaSerGlyProAlaGlyValArgGlyProAsnGlyAspAlaGlyArgProG1 442
Qy 641 -----GGAGCAGCTCGCCAGATT---CCTGG 663
Db 442 yGluProGlyLeuMetGlyProArgGlyLeuProGlySerProGlyAsnIleGlyProAl 462
Qy 664 GGGTGTCTGTGACAAGCCCGCAGCTGGAAGCCCT-----GACGAGCAGCTGCCACCAGC 717
Db 462 aGlyLys-----GluGlyProValGlyLeuProGlyIleAspGlyArgProGlyProI1 480
Qy 718 TGGT-----GGACCAGTGTGCA 735
```

```
Db 480 eGlyProValGlyAlaArgGlyGluProGlyAsnIleGlyPheProGlyProLys----- 498
Qy 736 ACGTGTAGGCGCTCCCGTGGCGGGAAGAGTTGGGCTGTGAAGACAT 787
Db 499 -----GlyProThrGlyAspProGlyLysAsn---GlyAspLysGlyHis 512
RESULT 7
US-10-196-749-525
; Sequence 525, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 525
; LENGTH: 4440
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-525
Alignment Scores:
Pred. No.: 0.0334 Length: 4440
Score: 115.00 Matches: 86
Percent Similarity: 73.3% Conservative: 50
Best Local Similarity: 23.6% Mismatches: 181
Query Match: 2.6% Indels: 48
DB: 6 Gaps: 10
US-10-768-158-1 (1-2419) x US-10-196-749-525 (1-4440)
Qy 58 GGGAGTTCGAGAGCAAGTACTTCGAGTTCATGGGTGGCGGTGGCGGCTTCTGCGCGC 117
Db 3325 GlyThrCysAlaAlaThrGlyThrCysCysThrCysThrThrCysAlaThrAlaThr 3344
Qy 118 GGAAGATCGAGAGATCCGCCAATTTCGGGTGGCGGCCAGGACGTGTGATCGTCACCT 177
Db 3345 GlyThrThrAlaAlaCysThrThrThrThrCysThrThrAlaThrThrGlyAlaThrThr 3364
```

```
Qy 178 ACCCCAGTCCGGCACCAGCTTGTCTGTCAGGAGGTGGTCTACTTGGTGAGCCAGCGCGCTG 237
Db 3365 ThrGlyThrThrAlaGlyAlaGlyCysAlaCysThrThrThrGlyThrAlaThrAlaThr 3384
Qy 238 ACCCCGATGAGATCGGCTTGATGAACATCG-----ACGAGCAGCTCCCGCTCCTGG 288
Db 3385 ThrGlyThrGlyAlaAlaThrAlaThrAlaGlyCysThrThrThrGlyThrGly 3404
Qy 289 AGTACCCACAGCCGGCTCGACATCATCAAGNACTGACCTCTCCCGGCTCATCAAGA 348
Db 3405 AlaThrCysAlaGlyAlaAlaThrThrAlaThrGly-----GlyAlaAlaThrGly 3422
Qy 349 GCCACCTGCCCTACCGCTTCTGCTCCTCTGACATCCACATGAGAGCTCCAAAGGTCTCT 408
Db 3423 ThrThrGlyThrThrGlyThrCys-----ThrThrThrAlaGlyThrThrThrGly 3440
Qy 409 ATATGGCTCGCAACCCCAAGGATCTGGTGGTGTCTTATTATCATGATTCACCGCTCTCTGC 468
Db 3441 CysThrThrAlaThr-----CysAlaAlaGlyThrThrThrAlaAlaAla 3455
Qy 469 GGACCATGAGTACCGAG-----GCACCTTCAAGATTTCTGCGGAGGTTTATGA 519
Db 3456 GlyCysCysAlaThrThrCysAlaGlyAlaGlyAlaThrGlyThrGlyThrAlaAla 3475
Qy 520 ATGATAAGCTGGGCTACGGCTCCTGTTGAGCAGCTGCAGGAGTCTCGGAGCACCGCA 579
Db 3476 AlaThrGlyThrGlyThrAlaThrGlyThrGlyThrThrAlaAlaAlaThrThrThr 3495
Qy 580 TGGACTCGAAACGTGCTTTTCTCAAGTATGAAGACATGCATCGGACCTGGTACGATGG 639
Db 3496 AlaThrGlyThrCysThrThrAlaThrThrThrGlyThrGlyThr----- 3511
Qy 640 TGGAGCAGCTGGCGAGATTCCTGGGGTGTCTGTGACAGGCCAGCTCGGAGCCCTGA 699
Db 3512 ThrThrThrCysThrGlyAlaThrGlyCysThrCysAlaThrAlaThrGlyThrThr 3531
Qy 700 CGGAGCACTGCCACGAGCTGGTGACCAAGTGTGCAACGCTGAGGCCCTCGCCGCGGCC 759
Db 3532 AlaGlyAlaAlaThrGlyGlyThrCysAlaAlaGlyCysAlaAlaThrCysCysAla 3551
Qy 760 GGGGAAGAGTTGGGCTGTGGAAGACATCTTCAACCGCTCCATGAATGAGAAGTTGACT 819
Db 3552 AlaCysThrThrAlaThrGlyAlaThrThrAlaCysAlaThrAlaAlaAlaThrAlaThr 3571
Qy 820 TGGTGTATAAACAGAACATGGGAAGTGTGACCTCAGCTCAAGTTTGACTTTTA----- 868
Db 3572 ThrCysThrCysCysCysAlaThrAlaThrThrThrGlyCysThr-CysThrThrAlaGl 3591
Qy 869 -----TTTATAATAACAGAAACAACTGTCATGCTCACAATACCAGACAGTCTAC 921
Db 3591 yCysAlaThrThrThrThrThrThrThrThrThrThrThrThrThrCysAlaThrGlyTh 3611
Qy 922 TAGCAAAAGTCTGTATGCATTCATTTATTTCTTGTGTCGACAAACT----- 968
Db 3611 rAlaAlaCysCysCysThrThrThrGlyThrThrCysThrThr-ThrCysAlaGlyGlyA 3631
Qy 969 -----CTGGAGACGCGTGTGAACACAGCGGGGAGGAGAGAGC----- 1007
Db 3631 laAlaThrThrAlaCysThrThrThrGlyCysGlyGlyThrAlaAlaGlyAlaAlaA 3651
Qy 1008 -----GGCGTGAGCGGAGGAGTGTGATGATTCACCAACCGAAGCAGCTGTCTCG 1056
Db 3651 laThrGlyGlyCysThrGlyCys-----ThrThrCysCysA 3665
Qy 1057 CCTTTAGAACG 1067
Db 3665 laGlyThrThr 3668
RESULT 8
US-10-985-570-1
; Sequence 1, Application US/10985570
; Publication No. US20060100138A1
; GENERAL INFORMATION:
```





```

Db 350 -----:||||:|||||-----lleGlyysProGly----- 354
Qy 251 CGGCTTGATGAACATCGACGAGCTCCCGTCTCT-----GGA 289
Db 355 -----PheProGlyProLysGlyAspArgGlyMetGlyGly 366
Qy 290 GTACCCACAGCC---GGGCTGGACATCATCAAGAACTGACCTCTCCCG--- 337
Db 367 ValProGlyAlaLeuGlyProArgGlyGluLysGlyProIleGlyAlaProGlyIleGly 386
Qy 338 ---CCTCATCAAGACCA---CCTGCCCTPACCGCTTCTGCCCTCTGACCTCCACAA 388
Db 387 GlyProProGlyGluProGlyLeuProGlyIlePro---GlyProMetGlyProPro--- 404
Qy 389 TGGAGACTCAAGGTCATCTATATGGCTGCGCAACCCCAAGGATCTGGTGGTCTCTTA--- 445
Db 405 -----GlyAlaIleGlyPheProGlyProLysGlyGluGlyGlyIleValGly 420
Qy 446 -----TTATCAGTTCCACCGCTCTCTGCCGAGCATGAGCTACCG 484
Db 421 ProGlnGlyProProGlyProLysGlyGluProGlyLeuGlnGlyPheProGlyLysPro 440
Qy 485 AGGCACCTTCAAGAACTCTCCGAGGTTTATGAATGATGAAGCTGAGCTACGCTCCTG 544
Db 441 GlyPheLeuGlyGluValGlyPro-----ProGlyMetArgGlyLeu 454
Qy 545 -----GTTTGACACGTCGACGAGTTCTGGGAGCACCCGACATGCGAAG 589
Db 455 ProGlyProIleGlyProLysGlyGluAlaGlyGlnLysGlyValPro---GlyLeu--- 472
Qy 590 CGTGCTTTTCTCAAGTATGAAGACATGCATCGGACCTCTGTGAGCATGGT-----GGA 643
Db 473 -----ProGlyValProGlyLeuGly 480
Qy 644 CGAGCTGCCAG-----ATTCTGGGGTGTCCTGTGACAAAGGCCA----- 685
Db 481 ProLysGlyGluProGlyIleProGlyIleProGlyAspGlnGlyLeuGlnGlyProProGlyIlePro 500
Qy 686 ---GCTGGAAGCCCTGACGAGCACTGCCACGAGCTGGTGGACCAAGTCTGCAACGCTGA 742
Db 501 GlyIleGlyGlyProSerGlyProIleGlyProProGlyIlePro----- 515
Qy 743 GGCCCTGCCCTGGGCGGGG 763
Db 516 GlyProLysGlyGluProGly 522

```

## RESULT 12

```

US-10-505-928-799
; Sequence 799, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; PRIOR FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 799
; LENGTH: 1384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-799

```

```

Alignment Scores:
Pred. No.: 0.131 Length: 1384
Score: 105.50 Matches: 109
Percent Similarity: 32.4% Conservative: 45
Best Local Similarity: 22.9% Mismatches: 174
Query Match: 2.4% Indels: 148

```

```

DB: 6 Gaps: 27
US-10-768-158-1 (1-2419) x US-10-505-928-799 (1-1384)
Qy 1140 CTTTATCTTACGGTCCAGGCCATTTGGAATCTGCAATGTGGAGACTGTT----- 1093
Db 103 LeuTyProCysGlyAsnAla-----CysThrIleAspGlyLeuProValArgGln 119
Qy 1092 TGTATATCAGACATGGAGAGGCTGCACGCTTCTAAAGCGGAGACAGCTGCTCGGT----- 1039
Db 120 ProThrArgLeuThrGlnGlyCysMetLeu-----CysLeuGlyGlnSer 134
Qy 1038 -----TGGGAATCATCATCATCTCCCTCCG 1015
Db 135 ThrPheLeuArgPheAsnHisProAlaGluAlaLysTrpMetLysSerMetIleProAla 154
Qy 1014 TCAGCGCCTCTCTCCCTCCCGCTGTT----- 985
Db 155 GlyGlyArgAlaProGlyProProTySerProValProAlaGluSerGluSerLeuVal 174
Qy 984 -----CACACGCTGCTTCCAGAGTTTGTCCAGCAAGGAATAAATGAATGCATACAG 934
Db 175 AsnGlyAsnHisThr-----ProGlnThrAlaThrArgGlyProSerAlaCysAlaSer 192
Qy 933 GACTTTTGGCTAGTACTGTCTGGGTATTGTGAGCATGCAGGTTGTTGTTTCTGTTATT 874
Db 193 HisSerSerLeuValSerSerIleGlyLysAspLeuGlnGluIleMetAspSerLeuVal 212
Qy 873 ATAAATAAAGTCAACGTCAGGTCACACTTTCCCATCTCTCTGTTTATACACCAAGTCAA 814
Db 213 LeuGluGluProGlyAlaAlaGlyLys-----LysProAlaAla 225
Qy 813 ACTTCTCATTCATCGAGACGGTGAAGATGTCCTTCCACAGCCCAACTCTTCCCC---GCG 757
Db 226 ThrSer-----ProLeuSer 230
Qy 756 CCAGCGGCAGGCTCAGCGTTGCAGCACTGGTCCACCAAGCTGGTGGCAGTCTCCGTC 697
Db 231 ProMetAlaAsnGlyGlyArgTyLeuLeuSerProPro-----ThrSerProGly 247
Qy 696 GGGCTTCAGCTGGGCTTGTGCAC---AGGACACCCCGAGGAATCTGCCAGCTGCTCCA 640
Db 248 AlaMetSerValGlySerSerTyLeuAsnThrSerProAlaPheSerProLeuSerSer 267
Qy 639 CCATCGTCACAGGTCCTCGATGTCATCTTTCATATCTTGAGAAAAAGCAGCTTCGAGTCCA 580
Db 268 ProAlaSerSerGlySerCysAlaSer---HisSerProSerGlyGlnGluProGlyPro 286
Qy 579 TGGCGTCTCCCAAACTCTCGACGCTCTCAAAACAGGACCGCTAGCCAGCTTATCAT 520
Db 287 SerValProProLeuValProAla-----ArgSerSerSerTyHis 300
Qy 519 -----TCATAAACCTCCGCGAGAATTTCTTGAAGGTGCTCGGT 481
Db 301 LeuAlaLeuGlnProProGlnSerArgProSerGlyAlaArgSerGlu----- 316
Qy 480 AGCTCATGGTCCGACAGAGCGGTGGAATCTGATTAATAAGACACCAACAGACTCTTGGGGT 421
Db 317 SerProArgLeuSerArgLysGlyGly-----HisGluArgProProSerPro-GlyLe 334
Qy 420 TGGCAGGCATATAGTACCTTGGAGTCTCATTGTG-----GAGGTCAAG----- 375
Db 334 uArgGlyLeuLeuThrAspSerProAlaAlaThrValLeuAlaGluAlaArgAlaTh 354
Qy 374 -----GGGCAGAAAGCGGTAGGCGAGGCTGCTCTTGTATAGGC--- 337
Db 354 rGluSerProArgLeuGlyGlyGlnLeuProValAlaIleSerLeuSerGluTyPr 374
Qy 336 -----GGGCAGAGGTCAAGTCTTGTGATGATGCCAGGC---CCGGCTGGGTACTCC 287
Db 374 oAlaSerGlyAlaLeuSerGlnProThrSerIleProGlySerProLysPheGlnProPr 394
Qy 286 AGGACCGGAGCTCTCGTCGATGTTTCATCAAGCCGATCTTCATCGGGGTGAGCCCTGG 227

```

```
Db 394 oValProAlaProArgAnlyIleGlyThrLeuGln-----AspArgProPr 410
Qy 226 CTCAACAGTAGACCACTCTCGCAGCAAGCTGGTGC----- 190
Db 410 oSerProPheArgGluProProGlySerGluArgValLeuThrThrSerProSerArgG1 430
Qy 189 -----CGACTTGGGGTAGTGACGATCCACAGTCGCTGGGCGGCACCGGAAG 140
Db 430 nLeuValGlyArgThrPheSerAspGlyLeuAlaThrArgThrLeuGlnProProGluSe 450
Qy 139 TTGGCGA-----TCTCTCCATCTCCCGCG 113
Db 450 rProArgLeuGlyArgArgGlyLeuAspSerMetArgGluLeuProProLeuSerProSe 470
Qy 112 C---AGAGGCGGCGCAGCGCCATGGAACTCGAAGTACT----- 73
Db 470 rLeuSerArgArgAlaLeuSerProLeuProThrArgThrThrProAspProLysLeuAs 490
Qy 72 -----TGTCTCGAACTCCC-----CCGGGTGCTGGGGTCTCGGCTCGCTCTCGCC 23
Db 490 nArgGluValAlaGluSerProArgProArgArgTrpAlaAlaHisGlyAlaSerProG1 510
Qy 22 A-----TGCGCGCGCGCTCGCGCTCGC 1
Db 510 uAspPheSerLeuThrLeuGlyAlaArgGlyArgArgThrArg 524

RESULT 13
US-10-559-415-190
; Sequence 190, Application US/10559415
; Publication No. US20060100132A1
; GENERAL INFORMATION:
; APPLICANT: Astrazeneca AB et al
; TITLE OF INVENTION: Diagnostic Method
; FILE REFERENCE: 101073-1P WO
; CURRENT APPLICATION NUMBER: US/10/559,415
; CURRENT FILING DATE: 2005-12-06
; PRIOR APPLICATION NUMBER: 0313081.2
; PRIOR FILING DATE: 2003-06-06
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 190
; LENGTH: 1809
; TYPE: PRT
; ORGANISM: Homo Sapiens - peptide sequence of amino acids
US-10-559-415-190

Alignment Scores:
Pred. No.: 0.194 Length: 1809
Score: 104.00 Matches: 200
Percent Similarity: 32.2% Conservative: 89
Best Local Similarity: 22.3% Mismatches: 309
Query Match: 6 Indels: 301
DB: Gaps: 46

US-10-768-158-1 (1-2419) x US-10-559-415-190 (1-1809)
Qy 2372 TGCAACGCCGGAACATGCTTTATAAAGATGCAAGCAAGCAGAGATCA----- 2319
Db 816 CysGlyValGlyGluAlaSerLeuAspLys-----AlaAspSerGluGlySerAsnSer 833
Qy 2318 -----TGTTAGTGTCTAGTAACATCCATACAGTACTATAAATAAGAAAAA 2274
Db 834 GlyGlyThrTrpProLysAlaMetLeuSerSerThrAlaValProGluLysLeuSerVal 853
Qy 2273 TATAAATCAAATTTACAGAAAGACGCTCCCTCAGAAACACAGGCAAGGTGTGCGCG 2214
Db 854 TyrLysLysProLysGlnArgLysSerIlePheAspProAsnThrPheLys----ArgPro 872
Qy 2213 AGCCACCATGCGATGCCACGTCGACCTGCTGGCCACCGCTTCTCCCATTTACAAACA 2154
Db 873 GlnThrProProLysIleAspTyrLeuLeuProGlyProGlyProAlaAlaHisSerProGln 892
```

```
Qy 2153 AARACACAAGCTTCTCTACACCGCGTTTCAAAATACAGCACCA----- 2112
Db 893 ProSerLys---ArgAlaGlyProLeuThrProProLysProProArgArgSerAspSer 911
Qy 2111 -----AGAACGGAGAGACCCCTTCC---ACAGCAGCGACCCCTCTCTGACCACC 2070
Db 912 IleLysPheGlnHisArgLeuGluThrSerSerGluSerGluAlaThrLeuValGlySer 931
Qy 2069 ACTTTGGGTGCTTTCAGCCACATATAATGACAGACACACCGCGTCCACCGGGTTGGCT 2010
Db 932 Ser-----ProSerThrSerProProSerAlaLeu--- 941
Qy 2009 GGCTACTGGCCCGC-----ACTCTCATCGCTTTCAGCGAATGGGG 1968
Db 942 -----ProProAspValAspProGlyGluProMetHisAlaSerProPro----- 956
Qy 1967 TGAGAAAAATAGGCAGTGAAGACCAATAAAGACCTCACCTTTACTCCCTCAGATACACGGAG 1908
Db 957 -----ArgLysAlaArgValArgIleAlaSerSerTyrTyrProGlu 970
Qy 1907 GGATAGGTGACAGACACGCG-----AGGGCGCGTCCCACTGATCCAGGTGAA 1857
Db 971 GlyAspGlyAspSerSerHisLeuProAlaLysLysSerCysAspGluAsp----- 987
Qy 1856 ACAGGGCTACACTCGCAAAATGGTCTCTCCACGCGCTCAGCGCTCAGGTGCACAGAACAA 1797
Db 988 -----LeuThrSerGlnLysValAspGluLeuGlyGlnLysArgArg----- 1001
Qy 1796 ACTGAATGATAAAGACCC-----GGGGATTTTAGAAATTTTAAAGCATCTGC 1746
Db 1002 -----ArgProLysSerAlaProSerPheArgProLysLeuAlaProValVal 1017
Qy 1745 TTGCACGANTTAGTAGCTTACTTCAGGAACAAAGTAAGCTGTGGATTTCAGCTCA 1686
Db 1018 IleProAla-----GlnPheLeuGluGlnLysCysValProAlaSerGly 1033
Qy 1685 AATTCCGAGCGCAACTCGGTCTTTGGTGGCG-----ACTCTCCAC 1644
Db 1034 GluLeuSerProGluLeuGlnGluTrpAlaProTyrSerProGlyHisSerSerArgHis 1053
Qy 1643 -----CCACTGTCATGCAAAATGCTCTGACTGTCATTAGACTCCACCTCCAGCGCTC 1593
Db 1054 SerAsnProProLeuTyrProSerArgProSer---ValGlyThrValProArgSerLeu 1072
Qy 1592 ACCAGGCAGACTTCCCTCAGCAC-ATGCTCGCGCGGCTCTCTCTGC----- 1549
Db 1073 ThrProSerThrThrValSerSerIleLeuArgAsnProIleTyrThrValArgSerHis 1092
Qy 1548 ---ATTGGGCTTTG---GAGCCGCCGCTGGA----- 1522
Db 1093 ArgValGlyProCysSerSerProProAlaAlaArgAspAlaGlyProGlnGlyLeuHis 1112
Qy 1521 ---AGTCTCTGGCAGAGAGCGTCCATTCTGAAGACCTAGTGAACAATTTCTGTCTCA--- 1468
Db 1113 ProSerValGlnHisGlnGlyArgLeuSerLeuAspLeuSerHisArgThrCysSerAsp 1132
Qy 1467 TTTCCAGGA----- 1459
Db 1133 TyrSerGluMetArgAlaThrHisGlySerAsnSerLeuProSerSerAlaArgLeuGly 1152
Qy 1459 ----- 1459
Db 1153 SerSerSerAsnLeuGlnPheLysAlaGluArgIleLysIleProSerThrProArgTyr 1172
Qy 1458 -----GGTTCTGAAGAAGTGAACCTCGTTCAGAAATAAACAATGTCAGAT 1414
Db 1173 ProArgSerValValGlySerGluArg---GlySerValSerHisSerGluCys-SerTh 1191
Qy 1413 TCCCCCAGAGAGACCCACTCATCTCTGTGTCTCACCACCTATCTCAGAACCAAGTGAGAA 1354
Db 1191 rProProGlnSerProLeuAsnIleAspThrLeuSerSerCysSerGlnSerGlnThrSe 1211
Qy 1353 CAGAAATGGCTGTGCTCTAGATA----- 1332
```



Db 1042 Ser-----ProSerThrSerProProSerAlaLeu--- 1051  
Qy 2009 GGCTACTGCCCCCG-----ACTCTCCATCGCTTCAGGAATGTGGG 1968  
Db 1052 -----ProProAspValAspProGlyGluProMetHisAlaSerProPro----- 1066  
Qy 1967 TGAGAAATAGGAGTGAAGCAATAAGACCTCACCTTTACTCCCTCAGATACACGAG 1908  
Db 1067 -----ArgLysAlaArgValArgLysAlaSerSerTyTyTyProGlu 1080  
Qy 1907 GGATGAGTGACAGACACG-----AGGGCGCGCTGCCAACCTGATCCGAGTAA 1857  
Db 1081 GlyAspGlyAspSerHisLeuProAlaLysLysSerCysAspGluAsp----- 1097  
Qy 1856 ACAGGGCTACACTCGCAAAATGGTCTCCCGCTGAGCGCTCAGCGTGCACAGAACAA 1797  
Db 1098 -----LeuThrSerGlnLysValAspGluLeuGlyGlnLysArgArg----- 1111  
Qy 1796 ACTGAATGATAAGACCC-----GGGATTTTAGAAATTTACTATTAAAGCATCTGC 1746  
Db 1112 -----ArgProLysSerAlaProSerPheArgProLysLeuAlaProValVal 1127  
Qy 1745 TTGCAGCATTTAGTGGCTATCACTTCAGGAAACAAAGTAAGGCTGTGGAATTCAGCTCA 1686  
Db 1128 IleProAla-----GlnPheLeuGluGlnLysCysValProAlaSerGly 1143  
Qy 1695 AATTCGAGCGCAACTCGTCTCTTTGGTGGCG-----ACTCTCCAC 1644  
Db 1144 GluLeuSerProGluLeuGlnGluTrpAlaProTySerProGlyHisSerSerArgHis 1163  
Qy 1643 -----CACTGCATGCAATGCTCTGACTGCATTAAGTACCTCCACCTGCCAGCTC 1593  
Db 1164 SerAsnProProLeuTyProSerArgProSer---ValGlyThrValProArgSerLeu 1182  
Qy 1592 ACCAGGCAGACTTCCTCCAGCAC-ATGCTCCGCGGGCTCCTCTGC----- 1549  
Db 1183 ThrProSerThrThrValSerSerIleLeuArgAsnProIleTyThrValArgSerHis 1202  
Qy 1548 ---ATTGGCCTTTG---GAGCGCGCCCTCGA----- 1522  
Db 1203 ArgValGlyProCysSerSerProProAlaAlaArgAspAlaGlyProGlnGlyLeuHis 1222  
Qy 1521 ---AGTCTCTGCAGAGGAGCGTCCATCTGAGACCTAGTGAACAAATTTCTGTCTCA--- 1468  
Db 1223 ProSerValGlnHisGlnGlyArgLeuSerLeuAspLeuSerHisArgThrCysSerAsp 1242  
Qy 1467 TTTCAAGGAA----- 1459  
Db 1243 TyrSerGluMetArgAlaThrHisGlySerAsnSerLeuProSerSerAlaArgLeuGly 1262  
Qy 1459 ----- 1459  
Db 1263 SerSerSerAsnLeuGlnPheLysAlaGluArgLysIleProSerThrProArgTy 1282  
Qy 1458 -----GGTCTCGAAGTGAAGTGAAGTCCGTCAGATATAAACAAGTCCAGAT 1414  
Db 1283 ProArgSerValValGlySerGluArg---GlySerValSerHisSerGluCys-SerTh 1301  
Qy 1413 TCCCCCAACAGACCAACTATCTCTGTTCTCCACCACTATCTCAGAACCAAGTGAGAA 1354  
Db 1301 rProProGlnSerProLeuAsnIleAspThrLeuSerSerCysSerGlnSerGlnThrSe 1321  
Qy 1353 CAGAAATGCTGTGCTTATAGTA----- 1332  
Db 1321 rAlaSerThrLeuProArgLysIleAlaValAsnProAlaSerLeuGlyGluArgLysAs 1341  
Qy 1331 -----ATAAACTACAAAATAAAGATTAAACCTTGA 1300  
Db 1341 pArgProTyValGluGluProArgHisValLysValGlnLys----- 1355  
Qy 1299 GATCTTTCTACCATTCGGGTGTGGCTCGCTCTGATTCCTCCCTTGGAAATGAACCTTTATT 1240  
Db 1240 ----- 1240

Db 1356 -----GlySer-----GluProLeuGlyIleSerIleValSe 1366  
Qy 1239 TGGTTTACTGACATTTATGTAGATTTCCAGTGAAGAGCTCTATAAAATACAAATATACGG 1180  
Db 1366 rGly----- 1367  
Qy 1179 GGTGAAAAGGACACACATTTCTAGTTGCATATATTACAGGCTTTTATCTTACGGTCCAGGC 1120  
Db 1368 -----GluLysGlyGlyIleTyValSerLysValThrValGlySerIleAlaHisGlnAl 1386  
Qy 1119 CATTTGGAACCTGCAATGTGGAGACTGTTTGTAAATCAGACATGGAGAGGCTGCACGTTCTAA 1060  
Db 1386 a---Gly-LeuGluTy-GlyAsp----- 1392  
Qy 1059 AGGCGACAGACTGCTCTCGGTGGGAATCATCACATCCCTCCGCTCAGCGGCTCTTCC 1000  
Db 1393 -----GlnLeuLeuGluPheAsnGlyIleAsnLeuArgSerAlaThrGluGlnGlnAl 1410  
Qy 999 CTTCCCCCGCTGTTTTCACACGCTGCTTCCAGAGTTCCTCCAGCAAGGAATAAATGAATGC 940  
Db 1410 laArgLeuIleIleGlyGlnGlnCysAspThrIleThrIleLeuAlaGlnTyArgProH 1430  
Qy 939 ATACAGGACTTTTGGCT-----AGTAGACTGTCTGGGTATTGTGACATGCAGGTGTT 886  
Db 1430 isValHisGlnLeuSerSerHisSerArgSerSerHisLeuAspProAlaGly-Thr 1449  
Qy 885 GTTCTCTGTTATTATAATAAAGTCAAACTGAGGTGCACACTTCCCATCT-----TCT 832  
Db 1450 HisSerThrLeu-----GlnGlySerGlyThrThrThrProGluHisProSer 1465  
Qy 831 GTTTTATACCAAGTCAAACTTCTCATTCATGAGAGCGGTGAAGATGCTCTCCACAGCC 772  
Db 1466 ValIleAspProLeuMetGluGlnAspGlu----- 1475  
Qy 771 CAACTCTTCCCGCCGCCAGCGGCTCAGCGTTCAGACACTGTGTCCACAGCTGTT 712  
Db 1476 -----GlyProSerThrProProAlaLysGlnSerSerArgIleAlaGly 1491  
Qy 711 GGCAGT-----GCTCCGTGAGGCTTCCA 688  
Db 1492 AspAlaAsnLysLysThrLeuGluProArgValValPheIleLysLysSerGlnLeuGlu 1511  
Qy 687 GCTGGGCT-----TGTCACAGGACACCCCGCAAGATCTGCCAGCT---GCTCCACCA 637  
Db 1512 LeuGlyValHisLeuCysGlyGlyAsnLeuHisGlyValPheValAlaGluValGluAsp 1531  
Qy 636 TCGTCAACA-----GTCCTCGATGTCATGT-----CTTCATCTTGTAG----- 600  
Db 1532 AspSerProAlaLysGlyProAspGlyLeuValProGlyAspLeu-IleLeuGluTyArg 1551  
Qy 599 -AAAAAGCAGCTTCGAGTCCATGCGGTGCTCCAGAACTCTCGACGTGTCAACACGAG 541  
Db 1551 ySerLeuAspValArgAsnLysThrValGluGluValTyValGluMetLeuLysProAr 1571  
Qy 540 AGCGGTACCGCAGCTTATCATTAACCTCCCGCAGAAATCTTGAAGGTGCTCGCT 481  
Db 1571 gAspGlyValArgLeuLysValGlnTyArgProGluGluPheThrLysAlaLysGlyLe 1591  
Qy 480 AGCTCATGTTCGCGAGAGCGGTGGAACCTGATATAAGACACACCAAGATCTTTGGGGT 421  
Db 1591 u-----ProGlyAspSerPheTyIleArgAlaLeuTyArgAspArgLeuAlaAspVa 1608  
Qy 420 TGCAGACCATATA-----GATGACTTGGAGTCTCCATTTGGAGGTGCAGAGG 373  
Db 1608 lGluGlnGluLeuSerPheLysLysAspAspIle---LeuTyValAspAspThrLeuPr 1627  
Qy 372 GCAGAAAGCGGTAGCGAGGTGCTCTTG-----ATGA 340  
Db 1627 oGlnGlyThrPheGlySer-TripMetAlaTrpGlnLeuAspGluAsnAlaGlnLysIleG 1647  
Qy 339 GCGGGGAGAGGTGAGTTCC-----TTGATGA 313  
Db 1647 lnArgGlyGlnIleProSerLysTyValMetAspGlnGluPheSerArgArgLeuSerM 1667



```
QY 312 TGTCCAGGCCCGCTGTGGTACTCCAGGACC-----GGGAGCTCTCGTCGA 265
      |||||
Db 1667 etSerGluValLysAspAsnSerAlaThrLysThrLeuSerAla-AlaAlaArgArg 1686
      |||||

QY 264 TGTTCATCAAGCCGATCTCATCGGGGTACGCGCCCTGGCTCA 223
      |||||
Db 1687 SerPhePheArgArgLysHisLysHisLysArgSerGlySer 1700
      |||||

RESULT 15
US-10-505-928-451
; Sequence 451, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 451
; LENGTH: 1821
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-451

Alignment Scores:
Pred. No.: 0.286 Length: 1821
Score: 102.00 Matches: 83
Percent Similarity: 33.1% Conservative: 22
Best Local Similarity: 26.2% Mismatches: 87
Query Match: 2.3% Indels: 125
DB: 6 Gaps: 18

US-10-768-158-1 (1-2419) x US-10-505-928-451 (1-1821)
QY 783 CTTTCCAGCCCAACTCTTCCCGGCCCCAGCGGAGG---CCTCAGCGTTGCAGCACT 727
Db 23 ProLeuThrLeuAlaLeuPheValGly-----AlaGlyHisAlaGlnArgAspProVal 40
      |||||

QY 726 GGT-----CCACCAGCTGGTGGCAGTGTCTCGTCCAGGCTTCCAGCTGGCGCTTGTCAC 673
      |||||
Db 41 GlyArgTyrGluProAlaGlyGlyAspAlaAsnArg-----Leu 53
      |||||

QY 672 AGGACACCCAGGAATCTGGCAGCTGCTCCACCATCG----- 634
      |||||
Db 54 ArgArgProGlyGlySerTyrProAlaAlaAlaAlaLysValTyrSerLeuPheArg 73
      |||||

QY 633 ---TCACCAGTCCGATGCTGCTTCTCATCTTGAGAAAGCAGCTTCGAGTCCATGC 577
      |||||
Db 74 GluGlnAspAlaProValAlaGlyLeuGlnProValGluArgAlaGlnProGly---Trp 92
      |||||

QY 576 GGTGTCCAGAACTCTCTGCAGCTGCTCAACACGAGGCGCTAGCCAGCTTATCATTC 517
      |||||
Db 93 GlySerProArgProThrGluAlaGluAlaArgProSerArgAlaGlnGlnSer 112
      |||||

QY 516 TAAA-----CCTCCGGGAGAAATCTTGAAGGTCCTCGGTAGTCTCAT----- 474
      |||||
Db 113 -ArgArgValGlnProProAlaGlnThrArgArgSerThrProLeuGlyGlnGlnGlnPr 132
      |||||

QY 473 -----GTCGCCGACAGCGCGT 457
      |||||

Db 132 oAlaProArgThrArgAlaAlaProAlaLeuProArgLeuGlyThrProGlnArgSer-- 151
      |||||

QY 456 GGAAGTATAAAGACACACCATCATCTTGGGGTTCCGAGCCATATAGATGACCTTGG 397
      |||||
Db 152 GlyAlaAlaProProThrProProArgGly-----ArgLeuThrGly 165
      |||||

QY 396 AGTCTCCATTGTGAGGTGACAGGCGCAAGACCGGTAGGCGAGGTGGCTCTTGATGAGGC 337
      |||||
```

```
Db 166 ArgAsnValCysGlyGlnCys----- 173
QY 336 GGGGAGAGGTTCAGTTCTTGATGATGTCACAGCCCGGCTGTGGGTACTCCAGGACCGGA 277
      |||||
Db 174 -----CysProGlyTrpThrThr----- 179
      |||||

QY 276 GCTGCTGCTCGATGTTTCATCAAGCGATCTCAT-----CGG 241
      |||||
Db 180 -----AlaAsnSerThrAsnHisCysIleLysProValCysGluProPro 194
      |||||

QY 240 GGTCAAGCCCTGCTGCTACCAAGTAGACCACTCTCTGCAGCAAGCTGTGCGGACTTGG 181
      |||||
Db 195 CysGlnAsnArgGlySerCysSerArgProGlnLeuCysVal-----CysArgSerGly 212
      |||||

QY 180 GGTAGGTGACCA----- 169
      |||||

Db 213 PheArgGlyAlaArgCysGluGluValIleProAspGluGluPheAspProGlnAsnSer 232
      |||||

QY 168 -----TCCACACGTCGCTGGCGCCACCGCGGAAGTTGGCGATCTCTCCATCTTCCCGC 115
      |||||
Db 233 ArgLeuAlaProArgArgTipAlaGlu-----ArgSerPro-----Asn 245
      |||||

QY 114 GGCAGAGCGCGCAGCGCCAGCCATCGAACT-----CGAAGTACTTGCTCT 67
      |||||
Db 246 LeuArgArgSerSerAlaAlaGlyGluGlyThrLeuAlaArgAlaGlnProProAlaPro 265
      |||||

QY 66 CGAACTCCCGCGGTGCTGGGGTCTCGGCTCTCGGCTCTCGCTCCGCCA 22
      |||||
Db 266 GlnSerProProAla-----ProGlnSerProPro 275
      |||||

RESULT 16
US-10-196-749-310
; Sequence 310, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
```

```
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-310

Alignment Scores:
Pred. No.: 0.203 Length: 596
Score: 101.50 Matches: 110
Percent Similarity: 36.7% Conservative: 49
Best Local Similarity: 25.4% Mismatches: 224
Query Match: 2.3% Indels: 50
DB: Gaps: 15

US-10-768-158-1 (1-2419) x US-10-196-749-310 (1-596)
Qy 61 AGTTCGAGAGCAAGTACT-----TCGAGTTCATGGCGTGCGGTGCGGC 105
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
159 SerSerArgAlaSerThrAlaThrAsnSerGluSerSerThrLeuSerSerGlyAla--- 177
Qy 106 CTTCTCGCGGGAAGATGAGGAGATCGCAACTTCCCGTGGCGGCCACGACGTGT 165
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
178 SerThrAlaThrAsnSerAspSerSerThrThrSerSerGlyAlaSerThrAlaThrAsn 197
Qy 166 GGATCGTCACCTACCCCAAGT-----CCGGCACCAGCTTCTGTCGAGGAGTGTCTACT 219
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
198 SerGluSerSerThrThrSerSerGlyAlaSerThrAlaThrAsnSerGluSerSerThr 217
Qy 220 TGGTGAGCCAGGGCGCTGACCCGATGAGATGCGCTTGATGAACATCGACGAGCTCC 279
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
218 ValSerSerArgAlaSerThrAlaThrAsnSerGlu---SerSerThrThrSerSerGly 236
Qy 280 CGTCTCGGAGTACCCACAGCGGGCGCTGGACATCATCAAGGAAGTACCTCTCCCGCC 339
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
237 AlaSerThrAlaThrAsnSerGluSerArgThrThrSerAsnGlyAlaGlyThrAlaThr 256
Qy 340 TCATCAAGAGCC-----ACCTGCCTACCGCTTCTGCCCTCTCACCTCCACAAATGGAG 393
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
257 AsnSerGluSerSerThrThrSerSerGlyAlaSerThrAlaThrAsnSerAspSerSer 276
Qy 394 ACTCCAAGGTCTATATGCTCGCAACCCCAAGGATCTGGTGGTGTCTATTATCATGAT 453
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
277 ThrValSerSerGlyAlaSerThrAlaThrAsnSer-----GluSer 290
Qy 454 TCCACCGCTCTCTGGGA---CCATGAGCTACCGAGGCACCTTTCAGAAATCTGCCGGA 510
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
291 SerThrThrSerSerGlyAlaSerThrAlaThrAsnSerGluSerSerThrThrSerSer 310
Qy 511 GGTTTATGAATGATAAGCTGGCTACGGCTCTCTGTTTGAGCAGCTGCAGGAGTCTGGG 570
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
311 Gly-----AlaSerThrAlaThrAsnSerAspSerSerThrThrSerSerGlyAla 327
Qy 571 AGCACCGCATGGACTCGAACGTGCTTTTCTCAAGATGTAAGACATGATCGGACCTGG 630
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
328 GlyThrAlaThrAsnSerGluSer-----SerThrValSerSerGlyIleSerThrVal 345
Qy 631 TGACGATGGTGAGCAGCTGGCCAGATTCCTGGGGGTGCTCTGTGACAGCCACGCTGG 690
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
346 ThrAsnSerGluSerSerThrProSerSerGlyAlaAlaSerThrAlaThrAsnSerGluSer 365
Qy 691 AAGCCCTCAGCGAGC-----ACTGCCACC-----AGCTGG 720
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
366 SerThrThrSerSerGlyAlaAsnThrAlaThrAsnSerGluSerSerThrValSerSer 385
Qy 721 TGGACCATGCTGCAACGCTGAGGCCCTGCGCGTGGCGGGGAGAGATTGGGTGTCGA 780
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
386 GlyAlaSerThrAlaThrAsnSerGluSerSerThrThrSerSerGlyValSerThrAla 405
Qy 781 AGGACATCTTACCGCTCTCCATGATGAGAAGTTTGACTTGGTGTATATAACAGAGATGG 840
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
406 ThrAsnSerGluSerSerThrThrSerSerGlyAlaSerThrAlaThrAsnSerAspSer 425
Qy 841 GAAAGTGTGACTCAGCTTACTTTTATTTATATAATAACAGAAACAACTGCTGCT 900
```

```
Db 426 SerThrThrSerSerGluAlaSerThrAlaThrAsn-SerGluSerSerThrValSerSe 445
Qy 901 CACAATACCCAGACAGTCTACTAGCCAAAGTCTCTGATGATTCATTTCCTTCTGCTG 960
Db 445 rGlyIleSerThrValThrAsnSerGluSerSerThrThrSerSerGlyAlaAsnThrAl 465
Qy 961 GACAAACTCTGGAAGCAGCGTGTGAAACAGCGGGGGAAGGAGAGCGGCTGAGCGGAG 1020
Db 465 aThrAsnSerGlySerSerVal-ThrSerAlaGlySerGlyThrAlaAlaLeuThrGlyM 485
Qy 1021 GGAGTGTGATGATTCCCAACCGAAGCAGCTGTCTCGCTTTAGAACGTGACGCTCTCCA 1080
Db 485 etHisThr-----ThrSerHisSerAlaSerThrAlav 496
Qy 1081 TGCTGTGATTACAAACAGCTCTCCACATTCGAGTTCATTCGCTGACCGTA---AGGATA 1137
Db 496 alSerGluAlaLysProGlyGlySer-LeuValPro-----TrpGluIlePheLeulle 513
Qy 1138 AAGCCTCTAATATATGCAACTAGAAATGCTGCGCTTTTCAACCCCGTATTATTGTTT 1197
Db 514 ThrLeuValSerValValAlaAlaValGlyLeuPheAlaGlyLeuPheCysValarg 533
Qy 1198 TAGAGCTTTTCACTGGAATCTACATAAATGTCAGTAAACCAATAAAGTTTCATTTC 1257
Db 534 AsnSerLeuSerLeuArgAsnThrPheAsnThrAlaValThrHisProHisGlyLeuAsn 553
Qy 1258 AGGGAATCAGGAGCGGACACACCCGCAATGGT 1290
Db 554 HisGlyLeu-----GlyProGlyProGlygly 562

RESULT 17
US-11-101-316-100
; Sequence 100, Application US/11101316
; Publication No. US20060099657A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
; FILE REFERENCE: P3230R1C17C1
; CURRENT APPLICATION NUMBER: US/11/101,316
; PRIOR FILING DATE: 2005-04-06
; PRIOR APPLICATION NUMBER: 10/063526
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 100
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-101-316-100

Alignment Scores:
Pred. No.: 0.203 Length: 596
Score: 101.50 Matches: 110
Percent Similarity: 36.7% Conservative: 49
Best Local Similarity: 25.4% Mismatches: 224
Query Match: 2.3% Indels: 50
DB: Gaps: 15
```

```
US-10-768-158-1 (1-2419) x US-11-101-316-100 (1-596)
Qy 61 AGTTCGAGCAAGTACT-----TCGAGTTCCATGGCGTGGCGTCCGCC 105
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
159 SerSerArgAlaSerThrAlaThrAsnSerGluSerSerThrLeuSerSerGlyAla--- 177
Qy 106 CTTCTCGCGGGAGATGAGGAGATCGCAACTCCCGGTGGCGGCCACGAGCTGT 165
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
178 SerThrAlaThrAsnSerAspSerSerThrThrSerSerGlyAlaSerThrAlaThrAsn 197
Qy 166 GGATCGTCACTACCCCAAGT-----CCGGCACCAGCTTCGTCAGGAGGTGCTACT 219
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
198 SerGluSerSerThrThrSerSerGlyAlaSerThrAlaThrAsnSerGluSerSerThr 217
Qy 220 TGGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCAGCAGCTCC 279
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
218 ValSerSerArgAlaSerThrAlaThrAsnSerGlu---SerSerThrThrSerSerGly 236
Qy 280 CGGTCTGGAGTACCCACAGCGGCGCTGGACATCATCAAGGAAGTACCTCTCCCGCC 339
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
237 AlaSerThrAlaThrAsnSerGluSerArgThrThrSerAsnGlyAlaGlyThrAlaThr 256
Qy 340 TCATCAAGAGC-----ACCTGCCCTACCGCTTCTGCCCTCGACCTCCACAAATGGAG 393
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
257 AsnSerGluSerSerThrThrSerSerGlyAlaSerThrAlaThrAsnSerAspSerSer 276
Qy 394 ACTCCAAGTCACTATATGCTCGCAACCCCAAGGATCTGGTGGTGTCTATTATCAGT 453
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
277 ThrValSerSerGlyAlaSerThrAlaThrAsnSer-----GluSer 290
Qy 454 TCCACCGCTCTCGGGA---CCATGAGCTACGAGGACACCTTCAAGAAATTCGCGGA 510
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
291 SerThrThrSerSerGlyAlaSerThrAlaThrAsnSerGluSerSerThrThrSerSer 310
Qy 511 GGTTTATGAATGAATAGCTGGCTACGCTCTCTGGTTGACGAGTGCAGGAGTCTGG 570
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
311 Gly-----AlaSerThrAlaThrAsnSerAspSerSerThrThrSerSerGlyAla 327
Qy 571 AGCACCGCATGGAATCGAAGCTGCTTTTCTCAAGTATGAAGACATCGACGAGCTGG 630
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
328 GlyThrAlaThrAsnSerGluSer-----SerThrValSerSerGlyIleSerThrVal 345
Qy 631 TGACATGGTGAGCAGCTGCCAGATTCCTGGGGGTCTCTGTACAGACCCAGCTGG 690
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
346 ThrAsnSerGluSerSerThrProSerSerGlyAlaAsnThrAlaThrAsnSerGluSer 365
Qy 691 AAGCCCTGACGGAGC-----ACTGCCACC-----AGCTGG 720
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
366 SerThrThrSerSerGlyAlaAsnThrAlaThrAsnSerGluSerSerThrValSerSer 385
Qy 721 TGGACCATGTGTCAACGCTGAGGCCCTCGCCCGTGGCGGGGAGAGTGGGCTGTGGA 780
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
386 GlyAlaSerThrAlaThrAsnSerGluSerSerThrThrSerSerGlyValSerThrAla 405
Qy 781 AGGACATCTTCACCGTCTCCATGAATGAGAAGTTTGACTTGGTGATAAACAGAGATGG 840
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
406 ThrAsnSerGluSerSerThrThrSerSerGlyAlaSerThrAlaThrAsnSerAspSer 425
Qy 841 GAAAGTGTGACCTCAGCTTTCACCTTTTATTAATACAGAAACACAACTGCATGCT 900
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
426 SerThrThrSerSerGluAlaSerThrAlaThrAsn-SerGluSerSerThrValSerSe 445
Qy 901 CACAATACCCAGACAGTCTACTAGCAAAAGTCTGTATGCTATTCATTTCCTTGCTG 960
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
445 rGlyIleSerThrValThrAsnSerGluSerSerThrThrSerSerGlyAlaAsnThrAl 465
Qy 961 GACAAATCTCGAGCAGCTGTGAACAGCGGGGAGGGAAGAGCGGCTGAGCGGAG 1020
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
465 aThrAsnSerGlySerSerVal-ThrSerAlaGlySerGlyThrAlaAlaLeuThrGlyM 485
Qy 1021 GGAGTGTGATGATCCCAACCGAGCAGCTCTCTCGCCTTTAGAACGTGACGCTCTCCA 1080
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
485 etHisThr-----ThrSerHisSerAlaSerThrAlaV 496
```

```
Qy 1081 TGCTGTGATTCAAAACAGTCTCCACATTGCAGTTCCAAATGGCTGGACCGTA---AGGATA 1137
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
496 alSerGluAlaLysProGlyGlySer-LeuValPro-----TrpGluIlePheLeuIle 513
Qy 1138 AAGCTGTATATATGCAACTAGAAATGTCTGCCCTTTTCAACCCCGTATTATTGTTT 1197
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
514 ThrLeuValSerValValAlaAlaValGlyLeuPheAlaGlyLeuPheCysValArg 533
Qy 1198 TAGAGCTTTTTCACGTGGAATCTACATAAATGTCAAGTAAACCAAAATAAAAGTTCA 1257
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
534 AsnSerLeuSerLeuArgAsnThrPheAsnThrAlaValThrHisProHisGlyLeuAsn 553
Qy 1258 AGGGAATCAGGAGCGAGCCACACCCGAATGGT 1290
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
554 HisGlyLeu-----GlyProGlyProGlyGly 562

RESULT 18
US-10-509-131-35
; Sequence 35, Application US/10509131
; Publication No. US20060089493A1
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec Inc.
; APPLICANT: Gately, Dennis
; APPLICANT: McLachlan, Karen
; TITLE OF INVENTION: NOVEL GENE TARGETS AND LIGANDS THAT BIND THERETO FOR TREATMENT AN
; TITLE OF INVENTION: DIAGNOSIS OF COLON CARCINOMAS
; FILE REFERENCE: 037003-0302886
; CURRENT APPLICATION NUMBER: US/10/509,131
; CURRENT FILING DATE: 2004-09-28
; PRIOR APPLICATION NUMBER: US 60/367,727
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/381,328
; PRIOR FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US 60/386,747
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 60/427,564
; PRIOR FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-509-131-35

Alignment Scores:
Pred. No.: 0.255 Length: 836
Score: 101.00 Matches: 166
Percent Similarity: 29.0% Conservative: 66
Best Local Similarity: 20.8% Mismatches: 264
Query Match: 2.3% Indels: 304
DB: Gaps: 39

US-10-768-158-1 (1-2419) x US-10-509-131-35 (1-836)
Qy 2345 AAGATGCAAGCAAGCAGACAGAGGATCATGTGTAGTCTCAGTAACATCCATACA----- 2292
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
170 LysSerLysGlyArgArgGlySerCysGlyAlaLeuAspThrLeuSerSerSerSer 189
Qy 2291 -----GTACTAAATAATAGAAAATATATAATTTTACAGAAAGCAGCCTC 2244
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
190 ThrSerAspCysAlaIleCysLeuGluLysIle----- 200
Qy 2243 CCTCACAGAAACACAGCAGAGGTGCTGCCGAGCCACCACCATGCGATGCCGTCACAC--- 2187
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
201 -----IleAspGlyGlyGluLeuArgValIleProCysThrHisArgPheHisArg 217
Qy 2186 -----CTGCCCTGGCAGCTTCTCCCATTTACAAACAAACAAACAC 2148
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
218 LysCysValAspProTrpLeuLeuGluHisThrCysProHisCysArgHis----- 235
Qy 2147 AAGCTTCTCTACACCCGTTTCAAATACAGCACCAAGACCGGAGACCCCTTCCACAGCA 2088
```



Publication No. US20060088532A1  
GENERAL INFORMATION:  
APPLICANT: Ludwig Institute for Cancer Research et al.  
TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES  
FILE REFERENCE: 28967/39178  
CURRENT APPLICATION NUMBER: US/10/505,928  
CURRENT FILING DATE: 2004-08-27  
PRIOR APPLICATION NUMBER: US 60/363,019  
PRIOR FILING DATE: 2002-03-07  
NUMBER OF SEQ ID NOS: 866  
SOFTWARE: PatentIn 3.2  
SEQ ID NO 831  
LENGTH: 2026  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-505-928-831

Alignment Scores:  
Pred. No.: 0.642 Length: 2026  
Score: 98.00 Matches: 102  
Percent Similarity: 30.8% Conservative: 35  
Best Local Similarity: 22.9% Mismatches: 144  
Query Match: 2.2% Indels: 164  
DB: 6 Gaps: 29

US-10-768-158-1 (1-2419) x US-10-505-928-831 (1-2026)

QY 37 CCGAGACCCCGAGCAGCGGGGAGTTCGAGAGCAAGTACTCTCGAGTTCATGGCGTGC 96  
DB 144 ProValProThrProProAspAlaCysArg-----GlyMetLeuCys 158  
QY 97 GGC-----TGCCGCCCTTCTCCCGGGGAGATGAGGAGATGCCCACTTCC 144  
DB 159 GlyPheGlyAlaValCysGluProAsnAlaGluGlyProGlyArgAlaSer----- 175  
QY 145 CGGTGCGGCCAGCAGCATGTGGATCGTCACCTACCCCAAGT---CCGGCACCAGTTCG 201  
DB 176 -----CysValCysLysLysSerProCysProSerValValAlaProValCys 191  
QY 202 TGCAGGAGTGGTCTACTTGGTGAGCCAGGGGGCTGACCCGATGAGATCGGCTTGATGA 261  
DB 192 GlySerAspAlaSerThyTyr---SerAsnGluCysGluLeuGlnArgAla----- 207  
QY 262 ACATCGACGAGCAGCTCCCGCTCTCGAGTACCCACAGCCGGGCTGACATCATCAAGG 321  
DB 208 ---GlnCysSerGlnGlnArgArgileArgLeuLeuSerArgGly----- 221  
QY 322 AACTGACCTCCCGCGCTCATCAAGAGCCACCTGCGCTTCTGCGCTTCTGCGCTTCTGACC 381  
DB 222 -----ProCysGlySerArgAspProCysSerAsnValThrCysSerPheGly 237  
QY 382 TCCACAATGGAGACTCCCAAGTTCATCTATATGGCTCGCAACCCCAAGGATCTGGTGTGT 441  
DB 238 SerThr-----CysAla-----Arg 242  
QY 442 CTTATTATCAGTTCACCGCTCTCTCGGAGCCATGAGCTACCGAGGCACCTTTCAAGAAT 501  
DB 240 ----- 501  
QY 502 TCTGCGGAGGTTTATGAATATAGTGGGTGAGTGGGTGAGTGGTGGTGGTGGTGGTGGT 555  
DB 243 SerAlaAspGlyLeu-----ThrAlaSerCysLeuCysProAlaThr 256  
QY 556 TGCAGGAGTCTGGGAGCAGCCATGGA----- 583  
DB 257 CysArg-----Gly-AlaProGluGlyThrValCysGlySerAspGlyAlaAspTyrPr 274  
QY 584 -----CTCGAAGCTGCTTTTCTCAA-----GTATGAAGACATGCG 618  
DB 274 oGlyGluCysGlnLeuLeuArgArgAlaCysAlaArgGlnGluAsnValPheLysLysPh 294  
QY 619 ATCGGAGCTGTGTACGATGTGTGGAGCAGCTGCGCAGATTCCTGGGGGTGCTCTGTGACA 678

Db 294 eAspGlyProCysAspPro-----CysGlnGlyAlaLeuPro----- 306  
QY 679 AGGCCAGCTGGAAGC-----CCTGACGAGGACACTGACC 711  
DB 307 -AspProSerArgSerCysArgValAsnProArgThrArgProGluMetArgLeuAr 326  
QY 712 ACCAGCTGGTGACAGCAGTGTGCAACGCTGAGGCGCTGCGCGTGGCGG----- 760  
DB 326 gProGluSerCysProAlaArgGln-----AlaProValCysGlyAspAspGlyValTh 344  
QY 761 -----GGGAAGAGTGGGCTGTGGAAGACATCTTACCCTGCT 798  
DB 344 rTyrGluAsnAspCysValMetGlyArgSerGlyAlaAlaArgGly-LeuLeuLeuGlnL 364  
QY 799 CCATGAATGAG-----AAGTTTGAAT 819  
DB 364 ysValArgSerGlyGlnCysGlnArgAspGlnCysProGluProCysArgPheAsnA 384  
QY 820 TGGTGTATTAACAGAAAGTGGAAAG-----TGTGAC---CTCAGCTTTGACT 864  
DB 384 laValCysLeuSerArgArgGlyArgProArgCysSerCysAspArgValThrCysAsp- 403  
QY 865 TTTATTTATATAACAGAAACACAACTGCTCATCTCACAA-----TACCCAG 912  
DB 404 -----GlyAlaTyrArgProValCysAlaGlnAspGlyArgThrTyrAspS 419  
QY 913 ACAGTCTACTAGCCAAAGTCTGTGTCATTCATTTATTTCTTCTGTCGACAACTCTGG 972  
DB 419 erAspCysTyrArgGlnGlnAlaGluCysArgGlnGlnArgAlaIleProSerLysHisG 439  
QY 973 AAGCAGCGTGTGAACAGCGGGGAGGAAGAGCGCGGTGAGC-----GGAGGGA 1023  
DB 439 lnGlyProCysAspGlnAlaProSerProCysLeuGlyValGlnCysAlaPheGlyAlaI 459  
QY 1024 GTGTGATGATTTCCCAAC---CGAAGCAGCTCTCTCGCTTTTGAAGAGTGCAGCTCTCCA 1080  
DB 459 hrCysAlaValLysAsnGlyGlnAlaIaCysGluCysLeuGlnAlaCysSer---SerL 478  
QY 1081 TGTCTGATTACAAACAGTCTCCACATTCGAGTTCGAATGCGCTGACCGTAAAGATAAAG 1140  
DB 478 euTyrAsp-----ProValCysGlySerAspGlyVal---ThrTyrGlySerA 493  
QY 1141 CCTGTATATATA 1151  
DB 493 laCysGluLeu 496

RESULT 20  
US-10-196-749-284  
Sequence 284, Application US/10196749  
Publication No. US20060094864A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C340  
CURRENT APPLICATION NUMBER: US/10/196,749  
CURRENT FILING DATE: 2002-07-16  
PRIOR APPLICATION NUMBER: 10/052586  
PRIOR FILING DATE: 2002-01-15  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059266  
PRIOR FILING DATE: 1997-09-18

```

; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 284
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-284

Alignment Scores:
Pred. No.: 0.316 Length: 258
Score: 97.50 Matches: 64
Percent Similarity: 28.5% Conservative: 17
Best Local Similarity: 22.5% Mismatches: 110
Query Match: 2.2% Indels: 93
DB: 6 Gaps: 13

US-10-768-158-1 (1-2419) x US-10-196-749-284 (1-258)
QY 816 CAACTTTTCATTCATGAGAGCGGTGACAGATGCTTCCACAGAGCCAACTCTTCCCGCG 757
DB 52 GluProSerHisProPheProargAlaThrAlaPro---ThralaGln-----Ala 67

QY 756 CCACGGGCGAGGCGCTCAGGGTTGACAGCTGTGCCACAGCTGGTGGCAGTCCCGTCA 697
DB 68 ProArgThrGlyProProargAlaThrValHisArgProLeuAlaAlaThrSerProAla 87

QY 696 GGGCTTCAGCTGGGCGCTTGTGCAGGACACCCAGGAATGTGCCAGCTGTCTCCACCA 637
DB 88 GlnSerPro-----GluThrThrPro---LeuTrpAlaThrAlaGlyPro 101

QY 636 TCGTCA-----CCAGGTCCGATGCGATGCTTCATCACTTGA 601
DB 102 SerSerThrThrPheGlnAlaProLeuGlyProSerPro----- 114

QY 600 GAAAGAACAGCTTCGAGTCCATGCGGTGCTCCAGAACTCTTCGACGTGCTCAACACAGG 541
DB 115 -----ThrThrProProAlaAlaGluArgThrSerThrThrSerGlnAlaPro 130

QY 540 AGCCGT---AGCCAGCTTATCATTCATAAACCTCCGGCAGAAATCTTGAAGGTGCCTC 484
DB 131 ThrArgProAlaProThrThrLeuSerThrThrThrGly----- 143

QY 483 GGTAGCTCATGTGCTCCGACAGAGCGGTGGAACCTGATAAGACACACACAGATCTTGG 424
DB 144 -----ProAlaProThrThrPro--- 149

QY 423 GGTGGGAGCCATATAGATGACCTTGGAGTCTCCATTGTGGAGGTGACAGGGGAGAAAGC 364
DB 149 ----- 149

QY 363 GGTAGGCGAGGTGGCTCTTGTATGAGCGGGGAGAGGTGATCTTGTGATGATGTCCAGGC 304
DB 150 -----ValAlaThrThrValProAla 156

QY 303 CCGGCTGTG---GGTACTCCAGGACGGGAGCTGCTCGTCGATGTTTCATCAAGCGCATCT 247
DB 157 ProThrThrProArgThrProThrProAspLeuProSerSerSerAsnSerSerValLeu 176

QY 246 CATCGGGGTGAGCGGCGCTGCACCAAGTAGACACCACTCTCTGCA---GCAAGCTGGTGC 190

```

```

; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 284
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-284

Alignment Scores:
Pred. No.: 0.316 Length: 258
Score: 97.50 Matches: 64
Percent Similarity: 28.5% Conservative: 17
Best Local Similarity: 22.5% Mismatches: 110
Query Match: 2.2% Indels: 93
DB: 6 Gaps: 13

US-10-768-158-1 (1-2419) x US-10-196-749-284 (1-258)
QY 816 CAACTTTCATTCATGAGAGCGTGAAGATGCTTCCACAGCCCACTCTTCCCGCG 757
DB 52 GluProSerHisProPheProargAlaThrAlaPro---ThralaGln-----Ala 67

QY 756 CCACGGCGAGGGCTCAGGGTTGACACTGTGTCACACAGCTGGTGGCAGTCCCGTCA 697
DB 68 ProArgThrGlyProProargAlaThrValHisArgProLeuAlaAlaThrSerProAla 87

QY 696 GGGCTTCCAGCTGGGCGCTTGTGCAGGACACCCAGGAATGTGCCAGCTGTCTCCACA 637
DB 88 GlnSerPro-----GluThrThrPro---LeuTrpAlaThrAlaGlyPro 101

QY 636 TCGTCA-----CCAGGTCCGATGTCATCTTCATCTTGA 601
DB 102 SerSerThrThrPheGlnAlaProLeuGlyProSerPro----- 114

QY 600 GAAAGACAGCTTTCAGTCCATGCGGTGCTCCAGAACTCTTCGACGTGCTCAACACAGG 541
DB 115 -----ThrThrProProAlaAlaGluArgThrSerThrThrSerGlnAlaPro 130

QY 540 AGCCGT---AGCCAGCTTATCATTCATAAACCTCGGCGAGAAATCTTGAAGGTGCCTC 484
DB 131 ThrArgProAlaProThrThrLeuSerThrThrThrGly----- 143

QY 483 GGTAGCTCATGTGTCGCCAGAGCGGTGGAACCTGATATAAGACACACACAGATCTTGG 424
DB 144 -----ProAlaProThrThrPro--- 149

QY 423 GGTGGCGGCAATATAGATGACCTTGGAGTCTCCATTGTGGAGGTGACAGGGCGAGAAAGC 364
DB 149 ----- 149

QY 363 GGTAGGCGAGGTGGCTCTTGTATGAGCGGGGAGAGGTGATCTTGTGATGATGTCCAGGC 304
DB 150 -----ValAlaThrThrValProAla 156

QY 303 CCGGCTGTG---GGTACTCCAGGACGGGAGCTGCTGCTCGATGTTTCATCAAGCGCATCT 247
DB 157 ProThrThrProArgThrProThrProAspLeuProSerSerSerAsnSerSerValLeu 176

QY 246 CATCGGGGTGAGCGGCTGGCTCACCAGTAGACACCACTCTCTGCA---GCAAGCTGGTGC 190

```

```
Qy 708 -----AGTGCTCCG-----TCAGGGCTTCAGCTGGCGCTTGTACACAGGACA 667
Db 149 ArgGluileSerAlaSerLeuLeuTyrglnAlaLeuProSerSerProAspHisGluSer 168
Qy 666 CCCCAGGAATCTGGCAGCTGCTCCACCATCGTCACAGGTCCTGCATGCTTCAT 607
Db 169 ProSerGlnGluSerProAsnAlaProThrSerThr----- 180
Qy 606 ACTTGAGAAAAAGCAGCTTCGAGTCCATCGCGTGTCCCGAGAACTCTCCGACAGTGTCTCAA 547
Db 181 -----AlaValLeuGlySerTrpGlySerProGlnProSerLeuAlaPro 196
Qy 546 ACCAGGACCTAGCCAGCTATCAT-----TCATAAACCTCCGGC----- 505
Db 197 ArgGluGlnGluAlaProGlyThrGlnTrpProLeuAspGluThrSerGlyIleTyTr 216
Qy 504 --AGAAATCTTGAAGGTGCTCGGTAGCTCA-----TGCTCCGCA 466
Db 217 ThrGluileArgGluArgGluArgGluLysSerGluLysGlyArgProProTrpAlaAla 236
Qy 465 GAGAGCGGTGAACCTGATAATAAGACACACACCATCATCT--TGCGGTTCGAGCCATAT 409
Db 237 ValValGly-----ThrProGlnAlaHisThrSerLeuGlnProHis 251
Qy 408 AGATGACCTTGAGTCTCCATTGTG----- 384
Db 252 HisHisProTrpGlu-ProSerValArgGluSerLeuCysSerThrTrpProTrpLysAs 271
Qy 383 -GAGGTACAGGGCAGAAAGCGGTAGGCGAGTGGCTTTCATGAGGCGGGAGAGGTCA 325
Db 271 nGluAspPheAsnGlnLysPheThr--GlnLeuLeuLeuLeuGlnArgProHisProAr 290
Qy 324 GTTCTTGATGATGCTCCAGGCGCGGTGTG-----GTACTCCAGGACCGGAGCT 274
Db 290 gSerGlnAspProLeuValLysArgSerTrpProAspTyTrValGluGluAsnArgGlyHi 310
Qy 273 GCTCGTCGATGTTTCATCAAGCGCATCTCAGCGGTGTCAGCGCCCTCGGTCTACCAAGTAGA 214
Db 310 sLeuileGluileArg-----AspLeuPheGlyProGlyLeuAspThrGlnGluProAr 328
Qy 213 CCACCTCCTCAGCAAGCTGTGCGGACTTGGG 180
Db 328 gileValileLeuGlnGlyAlaAlaGlyIleGly 339
```

## RESULT 22

US-10-511-937-2934

; Sequence 2934, Application US/10511937

; Publication No. US2006008836A1

## GENERAL INFORMATION:

; APPLICANT: EXPRESSION DIAGNOSTICS, INC.

; APPLICANT: Wohlgemuth, Jay

; APPLICANT: Fry, Kirk

; APPLICANT: Woodward, Robert

; APPLICANT: Ly, Ngoc

; APPLICANT: Prentice, James

; APPLICANT: Morris, MacDonald

; APPLICANT: Rosenberg, Steven

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING

; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION

; FILE REFERENCE: 506612000104

; CURRENT APPLICATION NUMBER: US/10/511.937

; CURRENT FILING DATE: 2004-10-19

; PRIORITY APPLICATION NUMBER: PCT/US2003/012946

; PRIORITY FILING DATE: 2003-04-24

; PRIORITY APPLICATION NUMBER: US 10/131,831

; PRIORITY FILING DATE: 2002-04-24

; PRIORITY APPLICATION NUMBER: US 10/325,899

; PRIORITY FILING DATE: 2002-12-20

; NUMBER OF SEQ ID NOS: 3117

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 2934

; LENGTH: 1429

; TYPE: PRT

; ORGANISM: Homo sapiens  
US-10-511-937-2934

## Alignment Scores:

Pred. No.:	0.617	Length:	1429
Score:	97.50	Matches:	70
Percent Similarity:	35.7%	Conservative:	27
Best Local Similarity:	25.7%	Mismatches:	105
Query Match:	2.2%	Indels:	70
DB:	6	Gaps:	15

US-10-768-158-1 (1-2419) x US-10-511-937-2934 (1-1429)

```
Qy 852 GGTACACATTTCCCATCTTCTGTTTATACACCAAGTCAAACTTCTCATTCATGAGACGG 793
Db 91 GlyHisSer--ProSerPheProTyTrSerProSerGlu--ProHisLeuGlySerPro 108
Qy 792 TGAAGATGTCCTTCCACAGCCCAACTCTTCCCGGCCACCGGAGGCGCTCAGGTTGC 733
Db 109 SerGlnProThrSerThrAlaValLeuMetProTrpIleHisGluLeuProAlaGlyCys 128
Qy 732 AGCACTGCT-----CCACCAGCTGGTGGC----- 709
Db 129 ThrGlnGlySerGluArgValLeuArgGlnLeuProAspThrSerGlyArgArgTrp 148
Qy 708 -----AGTGCTCCG-----TCAGGGCTTCCAGCTGGCGCTTGTACACAGGACA 667
Db 149 ArgGluileSerAlaSerLeuLeuTyrglnAlaLeuProSerSerProAspHisGluSer 168
Qy 666 CCCCAGGAATCTGGCAGCTGCTCCACCATCGTCACAGGTCCTGCATGCTTCAT 607
Db 169 ProSerGlnGluSerProAsnAlaProThrSerThr----- 180
Qy 606 ACTTGAGAAAAAGCAGCTTCGAGTCCATCGCGTGTCCCGAGAACTCTCCGACAGTGTCTCAA 547
Db 181 -----AlaValLeuGlySerTrpGlySerProGlnProSerLeuAlaPro 196
Qy 546 ACCAGGACCGTACGCCAGCTTATCAT-----TCATAAACCTCCGGC----- 505
Db 197 ArgGluGlnGluAlaProGlyThrGlnTrpProLeuAspGluThrSerGlyIleTyTr 216
Qy 504 --AGAAATCTTGAAGGTGCTCGGTAGCTCA-----TGCTCCGCA 466
Db 217 ThrGluileArgGluArgGluArgGluLysSerGluLysGlyArgProProTrpAlaAla 236
Qy 465 GAGAGCGGTGAACCTGATAATAAGACACACCATCATCT--TGCGGTTCGAGCCATAT 409
Db 237 ValValGly-----ThrProGlnAlaHisThrSerLeuGlnProHis 251
Qy 408 AGATGACCTTGAGTCTCCATTGTG----- 384
Db 252 HisHisProTrpGlu-ProSerValArgGluSerLeuCysSerThrTrpProTrpLysAs 271
Qy 383 -GAGGTACAGGGCAGAAAGCGGTAGGCGAGTGGCTTTCATGAGGCGGGAGAGGTCA 325
Db 271 nGluAspPheAsnGlnLysPheThr--GlnLeuLeuLeuLeuGlnArgProHisProAr 290
Qy 324 GTTCTTGATGATGCTCCAGGCGCGGTGTG-----GTACTCCAGGACCGGAGCT 274
Db 290 gSerGlnAspProLeuValLysArgSerTrpProAspTyTrValGluGluAsnArgGlyHi 310
Qy 273 GCTCGTCGATGTTTCATCAAGCGCATCTCAGCGGTGTCAGCGCCCTCGGTCTACCAAGTAGA 214
Db 310 sLeuileGluileArg-----AspLeuPheGlyProGlyLeuAspThrGlnGluProAr 328
Qy 213 CCACCTCCTCAGCAAGCTGTGCGGACTTGGG 180
Db 328 gileValileLeuGlnGlyAlaAlaGlyIleGly 339
```

## RESULT 23

US-11-238-282-19

; Sequence 19, Application US/11238282

; Publication No. US20060089303A1





```
Qy 1568 ATGCTGCTGAGGAAGTCTG-----CTGGTGGAGCTGGCAGGTGGAGTC 1612
Db 463 rCyProGluAenHisTyValGlnThrGlyHisProCyGlyLeuAenGlnTrp----- 481
Qy 1613 TAATGCAGTCAGGAGCATTTGCATGTCAGTGGTGGAGAGTCGGCCACCAAGGACCGAGT 1672
Db 482 -----lleCyHisIle-AspGly----- 486
Qy 1673 TGCCTCGGAATTGAGCTGAATTCACAGCCTTACTTTGTTCTGTAAGTGATAGCCTA 1732
Db 487 -----ValCyMet----- 489
Qy 1733 CTAATGCTGGCAAGCAGATGCTTAATAGTAATTTCTAAATCCCGGGTCTTTATCATTT 1792
Db 490 -----SerGlyAspGlnCysThrAspThrPheGlyLysGluValGluPheGlyProS 508
Qy 1793 CAGTTGTTCTGTCGACCTCAGG--CGCTCAGCGGTGGAGGA----- 1833
Db 508 erGluCysTySerHisLeuAenSerLysThrAspValSerGlyAsnCyGlyIleSerA 528
Qy 1834 -----CAATTTTGGAGTGTAGCCCTGTTCTACCTGGATCAGGT-----T 1873
Db 528 spSerGlyTyThrGlnCyGlyAlaAseAsnLeuGlnCyGlyLysLeuIleCysLysT 548
Qy 1874 GGCAGCGCGCTGGGTCTGTCCACCTCATCCTCGGTATCTGAGGAGTAAAGGT 1933
Db 548 yrValGlyLysPheLeuLeuGlnIleProArgAlaThrIleTyAlaAsnIleSer- 567
Qy 1934 GAGTCTTTATGTTCTCACTGCCCTAAATTTCT-----CACCCATTTGCTGGAAGCAT 1987
Db 568 --GlyHisLeuCyHisIleAlaValGluPheAlaSerAspHisAlaAspSerGlnLysMetT 587
Qy 1988 GGAGAGTCGGGGCGCAGTAGCAGCAACCCGCTGGGAGCGGGTGTCTGTCTCATTTAT 2047
Db 587 rpIleLysAspGly-----ThrSerCysGlySerAsnLysValCys----- 600
Qy 2048 GTGCTGGAAGCACCACCAAGTGTGTGTCAGGAGGTGCTGTGTGAAGGGTCTCCG 2107
Db 601 -----ArgAsnGlnArgCysValSerSer---SerT 610
Qy 2108 TTCTTGTGTGTATTGMAACGGGTGTA-----GAGAGAAGTGTGT----- 2151
Db 610 yrLeuGlyTyAspCysThrThrAspLysCysAsnAspArgGlyValCysAsnAsnLysL 630
Qy 2152 -----TTTTTGTCTAATGGGAGAGCGT 2176
Db 630 ysHisCysHisCysSerAlaSerTyLeuProProAspCysSerValGlnSerAspLeuT 650
Qy 2177 GGCAGGAGGTGGCAGTCGTCATCGCATCGCATGGTGGCTC-----GGCAGCACCTTGGCTG 2230
Db 650 rpProGly-----GlySerIleAspSerGlyAsnPheProV 663
Qy 2231 TGTTTCTGTGAGGAGGTCTTCTGTGAAATTTCAATTTAT----- 2272
Db 663 alAlaIleProAlaArgLeuProGluArg-TyrIleGluAenIleTyHisSerLys 682
Qy 2273 -----ATTTTTCTATTTTACTGTATGGATGTTACTGAGCTACAC 2317
Db 683 ProMetArgTrpProPhePheLeuPheIleProPhePhe----- 695
Qy 2318 ATGATCTCTGTGC---TTGCTTGATCTTTAATAAGACATGTTCCCGCGGTGCAAA 2374
Db 696 ---IleIlePheCysValLeuIleAlaIleMetValLysValaAsnPheGlnArgLysLys 714
RESULT 24
US-10-505-928-50
; Sequence 50, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505.928
```

```
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 50
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: KIAA0582
US-10-505-928-50

Alignment Scores:
Pred. No.: 0.584 Length: 760
Score: 96.50 Matches: 143
Percent Similarity: 32.5% Conservative: 60
Best Local Similarity: 22.9% Mismatches: 228
Query Match: 2.2% Indels: 194
DB: 6 Gaps: 30

US-10-768-158-1 (1-2419) x US-10-505-928-50 (1-760)
Qy 1702 CTGTGGAATTCAGCTCAATTCGAGCGCAACTCGTCTTTGGTGGCGACTCTCCACC 1643
Db 58 ValTrpGlyAlaGluGlyIleProAlaProThrCysTrpIleGly----- 72
Qy 1642 CACTGCATGCAATTCCTCTGACTGCATTAGACTCCCACTGCCAGCCTCACCAGGCAGA 1583
Db 73 -----ThrAspProGlyGly 77
Qy 1582 CTTCCTCCACACATGCTCCGCGGGT-----CTCTGCATTTGGGCC 1541
Db 78 ProSerArgAlaHisGlnProGlnAlaSerAspAlaAsnArgGluProVal-AlaGluAr 97
Qy 1540 TTGGAGCGCGCTCGAAGTCTCTGGCAGAGGAGCTCCATTTCTGAAGACCTAGTGAA 1481
Db 97 gSerGluProAlaLeuSerGlyLeuProProAlaThrMetGlySerGlyAspLeuLeu-- 116
Qy 1480 CAATTTCTGCTCATTTTCAGGAAGTCTGGAAGTGAAGTCCCTCAGATAAACAAGTG 1421
Db 117 -----LeuSerGlyGluSerGlnValGluLysThrLysLeuSerSerSe 131
Qy 1420 TCCAGATTCCTCCCAACAGACCCAACTC-----ATCCTCTGTTCTACCACTA 1373
Db 131 rGluGluPheProGlnThrLeuSerLeuProProArgThrThrIleCysSerGlyHisAs 151
Qy 1372 TCTCAGAACCAAGTGAACAGATGCTGT-----GCCTTAGATAATAAAC 1325
Db 151 pAlaAspThrGluAspAspProSerLeuAlaAspLeuProGlnAlaLeuAspLeuSerGl 171
Qy 1324 TACAAAAATAAGAGTTAACCTGAGATCTTTTCTTACCATTTGGGTGGCTCGCTCTGA 1265
Db 171 nGlnProHisSerSer-----GlyLeuSerCysLeuSerGlnTrpLysSerValLe 188
Qy 1264 TTCCCTTGGAAATGAATTTTATTTGTTTACTGACATTTATGATATTATTCAGTGAAA 1205
Db 188 uSerPro----- 190
Qy 1204 AGCTCTATAAATACATAATACGGGGTTGAAAAGGCAGACATTTCTAGTTGCATATATTA 1145
Db 191 -----GlySerAlaAlaGlnProSerSerCys-SerIle- 201
Qy 1144 CAGGCTTTATCTTACGGTCCAGGCCATTGCAATGTGGAGACTGTTTGTGAATCA 1085
Db 202 -----SerAlaSerSerThrGlySerSerLeuGln-----G 212
Qy 1084 GACATGGAGAGGTGCAGCTTTAAAGCGCAGACAGCTGTTCCGTTGGGAATCATCACA 1025
Db 212 lyHisGlnGluArgAlaGluProArgGlyGlySerLeuAlaLysVal---SerSer-Ser 230
Qy 1024 CTCCCTCCGCTCAGCGCGCTCTTCCCTTCCCGCTGTGTTCA-----CACGCTGCT 974
```

```
Db 231 LeuGluProValValProGlnGluProSerSerValValGlyLeuGlyProArgProGln 250
Qy 973 TCCAGAGTTGTCAGCAAGAAATAAATGAATGATACAGGACTTTTGGCTAGTAGACTG 914
Db 251 TrpSerProGlnProValPheSerGlyGlyAspAlaSerGlyLeuGlyArgArgLeu 270
Qy 913 TCTGGGTATTGTGAGCATGCGAGTTGTTCTTCTGTTATTATAATAAAAGTCAACGCTG 854
Db 271 SerPheGlnAlaGluTyrTrpAlaCys-Val- 280
Qy 853 AGGTACACACTTCCCATCTCTCTGTTTATACACCAAGTCAAACTTCTCATTTCA- 800
Db 281 -----LeuProAspSerLeu---ProProSerProAspArgHisSerProLeuTr 296
Qy 799 GAGACGTGAAGATGTCCTTCC- 767
Db 296 pAsn-----ProAsnLysGluTyrGluAspLeuLeuAspTyrThrTyrProLe 312
Qy 766 CTTCCCGCGCCAGCGGCGCTCAGCGTTGTCAGCACTGGTCCACCGCT----- 715
Db 312 uArgProGlyProGlnLeu---ProLysHisLeuAspSerArgValProAlaAspProVa 331
Qy 714 -----GGTGGCAGTGTCCGTTCAGGCTTCCAGCTGGGCTTGTTCAC-----AG 671
Db 331 lLeuGlnAspSerGlyValAspLeuAspSerPheSerValSerProAlaSerThrLeuLy 351
Qy 670 GACACCCCGAGAAATCTGGCCAGCTGCTCCACCATCGTCACAGGTCCTCGATGATGTCT 611
Db 351 sSerProThrAsnValSerProAsnCysProAla----- 363
Qy 610 TCATACTTGAGAAAGACAGCTTCAGAGTCCAGTCCGCTCCAGACCTCCGTCACGTCG 551
Db 364 -----GluAlaThrAlaLeuProPheSerGlyProArgGluProSerLeuLy 379
Qy 550 TCAAAACAGGAGCGGTAGCCAGCTTATCATTTATAAACTCCGCGCAAACTTCTTGAAG 491
Db 379 sGlnTrpProSerArgValProGlnLysGln-----GlyGlyMetGlyLeu---- 394
Qy 490 GTGCTCGGTAGTCATGGTCCGAGAGAGCGGTGGAACTGATATAAGACACACCCAGA 431
Db 395 -----AlaSerTrpSerGlnLeuAla-----SerThrPro----- 404
Qy 430 TCCTTGGGGTTCGAGCCATATAGATGACCTTGGAGTCTCCATTTGGAGGTACAGGGC 371
Db 405 -----ArgAlaProGlySerArgAspAlaArgTrpGluArgAr 417
Qy 370 AGAAAGCGGTAGGCGAGTGGCTCTTGATAGGCGGGAGAGGTGAGTTCCTTGTATGATG 311
Db 417 gGluProAlaLeuArgGly-AlaLysAspArgLeuThrIleGlyLysHisLeuAsp-Met 436
Qy 310 TCCAGGCGCGGTGTGGTACTCCAGGACCGGAGCTGCTCGTCGATGTTTCATCAAGCG 251
Db 437 Gly-SerProGlnLeuArgThrArgAspArgGly-----TrpProSerProAr 452
Qy 250 ATCTCATCGGGTCAGCGCCCTGGCTCCACCAAGTAGACACCTCTCGCAGCAAGTGTG 191
Db 452 gProGluArgGluLysArgThrSerGlnSerAlaArgProThrCysThrGlu----- 470
Qy 190 CCGGACTTGGGGTA-----GGTACGATCCACACGTCGCTGGCGCGC 149
Db 471 -SerArgTrp-LysSerGluGluGluValGluSerAspAspGluTrpLeuAlaLeuProA 490
Qy 148 ACCGG-----GAAGTTGGCGATCTCTCC 125
Db 490 laArgLeuThrGlnValSerSerLeuValSerTyrLeuGlySerIleSerThrLeuVal 510
Qy 124 ATCTTCCCGCGCAGAA-----GGCGCGCAGCGCCACCGCATGAACTCGAAGTACTTG 71
Db 510 hrLeuProGlyAspAlaLysGlyGlnSerPro-----LeuGluValSerA 526
Qy 70 CTCTCGAACTCCCGCGGGGTGTGGGGTCTCGGCTCGCTCTCCGCTATGCCGCGC 14
Db 526 sp-SerAspGlyProAlaSerPheProSerSerSerGlnSerGlnLeuProPro 544
```

## RESULT 25

```
US-10-505-928-227
; Sequence 227, Application US/10505928
; Publication No. US2006008532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178 US/10/505,928
; CURRENT APPLICATION NUMBER: 2004-08-27
; PRIOR FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 227
; LENGTH: 1663
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-227

Alignment Scores:
Pred. No.: 0.793 Length: 1663
Score: 96.50 Matches: 100
Percent Similarity: 35.5% Conservative: 70
Best Local Similarity: 20.9% Mismatches: 174
Query Match: 2.2% Indels: 135
DB: 6 Gaps: 23

US-10-768-158-1 (1-2419) x US-10-505-928-227 (1-1663)
Qy 1118 AATAAAGAGTTAAACCTGAGATCTTTTACCATTCGGGTGGCTCGCTCTGATTCCCC 1259
Db 329 AsnLysGluLeuPheLeuGlnSerHisThrGluIleGlyValSer----- 343
Qy 1258 TTGGAATGAACATTTTATTGTTTACTGACATTTATGTAGATTTCCAGTGAAGCTCT 1199
Db 344 -----TyrGlnTyrAlaLeuAspLeuGln 351
Qy 1198 ATAATAATACATAATACGGGTTCGAAAGGAGACATCTTAGTTGCATATATTACAGCT 1139
Db 352 ThrGlnHisAsnHisPheAlaMetAsnSerMetAsnAlaTyrValAsnIleAsnArgIle 371
Qy 1138 TTATCCTTACGTCAGGCCATTTGGAACGTG-----CAA 1106
Db 372 MetSerValAlaSerArgLeuSerGluAlaGlyHisTyrAlaSerGlnGlnIleLysGln 391
Qy 1105 TGTGAGACACTGTTTGTAAATCAGACATGGAGAGGCTGCACGTTCTAAAGCGGAGACAGCTG 1046
Db 392 -IleSerThrGlnLeuAspGlnGluTrpLys----- 401
Qy 1045 CTTCCGTTGGGAATCATCACACTCCCTCCGCTACCGCGCTCTCCCTTCCCGCTGTT 986
Db 402 ---SerPheAlaAlaLeuAspGluArgSerThrIleLeuAlaMetSerAlaValph 420
Qy 985 TCACACGCTGCTCCAGAGTTGTCCAGCAAGGAATAAATGAATGCATACAGGACTTTTG 926
Db 420 eHisGlnLysAlaGluGlnPheLeu---SerGlyVal-----AspAlaTr 434
Qy 925 GCTAGTAGACTGCTCTG-----GGTATTGTGAGCATGCAGGTGTTGTTTCTGTTATTAT 872
Db 434 pCysLysMetCysSerGluGlyGlyLeuProSerGluMetGlnAspLeuGluLeuAla 454
Qy 871 AAATAAAGTCAACAGTGAGGTGACACTTTCCCATCTCTCTTTAT-----ACACC 821
Db 454 eHisHisGlnThr-----LeuTyrGluGlnValThrGl 466
Qy 820 AAGTCAAACTTCTCATTCATTCATGCAGACGCTGAA-----GATGTCCTTCCACAGCCCA 770
Db 466 nAlaTyrThrGlu-ValSerGlnAspGlyLysAlaLeuLeuAspValLeuGlnArgPro- 485
Qy 769 ACTCTTCCCGCGCGGCGGCGGCGCTTCAGCGTTGCAGCA-----CTGGTCC 722
```

```
Db 486 --LeuSerProGlyAsnSerGluSerLeuThrAlaThrAlaAsnTyrSerLysAlaValH 505
Qy 721 ACCAGCTGGTGGCAGTGTCCGCTCAGGCTTCCAGCTGGCTTGTACAGGACACCC 662
Db 505 isGlnValLeuAspValValHis-----512
Qy 661 AGGAATCTGGCCAGCTGTCCACCATCTCACCAGTCCGATGCTTCTACTTG 602
Db 513 -----GluValLeuHisGlnArgLeuGluSer-----1letpg 526
Qy 601 AGAAAAAGCAGCTTCCAGTCCATGCTGCCAGAACTCTCCAGCTGTCAAAC 542
Db 526 inHisArgLysValArgLeuHisGlnArgLeuGlnLeuCysValPheGlnGlnAspValG 546
Qy 541 GAGCGGTAGCCAGCTTATCATTCATAAACCCTCCGCGCAGAAATCTTGAAGGTGCTCG 482
Db 546 inGlnValLeuAspTrpIleGluAsnHisGlyGluAlaPheLeuSerLysHisThrGlyV 566
Qy 481 TAGCTCATGTCGCGAGAGCGGTGGAACTGATTAAGACACACAGATCTCTGGGG 422
Db 566 alGlyLysSerLeuHisArgAlaArgAlaLeuGlnLysArgHisAspAspPheGluGluV 586
Qy 421 TTGGAGCATATAGATGACCTTGGAGTCTCCATTGTGGAGTCCAGAGGGCAGAAAGCG- 363
Db 586 alAlaGlnAsnThrTyrThrAsnAlaAspLysLeuLeuGluAlaAlaGluGlnLeuAlaG 606
Qy 362 --GTAGGAGCGGTGGCTTGTATGAGCGGGGAGAGGT-----327
Db 606 inThrGlyGluCysAspProGluGluIleTyrLysAlaAlaArgHisLeuGluValArgI 626
Qy 326 ----CAGTTCTTGATGATGTCACGGCCCG-----300
Db 626 leGlnAspPheValArgValGluGluGlnArgLysLeuLeuLeuAspMetSerValSerP 646
Qy 299 -----CTGTGGGA-----CTCCAGGAC-----CGGAGCTGCTCG 269
Db 646 heHisThrHisThrLysGluLeuIleThrTrpMetGluAspLeuGlnLysGluMetLeuG 666
Qy 268 TCGATGTTTCATCAAGCCGATCTATCGGGGTGAGCGCCCTGGCTCACCAGTAGACACC 209
Db 666 luAspVal---CysAlaAsp-----SerValAspAlaValGlnGluLeuLysGlnP 683
Qy 208 TCCTGACAGCTGTGTCGGAGCTTGGGTA-----GGTAGC 170
Db 683 heGlnGlnGlnThrAlaThrLeuAspAlaThrLeuAsnValIleLysGluGlyGluA 703
Qy 169 AT-----CCACAGCTGCTGGCGCCGACCGGGAAGTTGGCG 134
Db 703 spLeuIleGlnGlnLeuArgSerAlaProProSerLeuGlyGluProSerGluAlaArgA 723
Qy 133 ATCTCTCTCATCTTCCCGCGCAGAAAGCGCGCAGCGCAGCCATGGAACCTC 81
Db 723 sp-----SerAlaValSerAsnAsnLysThrProHisSerSerSerile 737

RESULT 26
US-11-140-487A-772
; Sequence 772, Application US/11140487A
; Publication No. US20060093617A1
; GENERAL INFORMATION:
; APPLICANT: Imogenetics N.V.
; TITLE OF INVENTION: Peptides for inducing a CTL and/or HTL response to Hepatitis C vi
; FILE REFERENCE: 166
; CURRENT FILING DATE: 2005-05-31
; PRIOR APPLICATION NUMBER: US/11/140,487A
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: EP 04012951.2
; PRIOR FILING DATE: 2004-10-28
; PRIOR APPLICATION NUMBER: EP 04447239.7
; PRIOR FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: EP 05102441.2
; PRIOR FILING DATE: 2004-06-03
; PRIOR APPLICATION NUMBER: US 60/576,310
; PRIOR APPLICATION NUMBER: US 60/622,782
```

```
; PRIOR FILING DATE: 2004-10-29
; PRIOR APPLICATION NUMBER: US 60/665,395
; PRIOR FILING DATE: 2005-03-25
; NUMBER OF SEQ ID NOS: 2278
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 772
; LENGTH: 3020
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-11-140-487A-772

Alignment Scores:
Pred. No.: 1 Length: 3020
Score: 96.50 Matches: 118
Percent Similarity: 31.7% Conservative: 48
Best Local Similarity: 22.5% Mismatches: 156
Query Match: 2.2% Indels: 203
DB: 7 Gaps: 28

US-10-768-158-1 (1-2419) x US-11-140-487A-772 (1-3020)
Qy 1175 GAAAGGCGAGCATCTAGTTGCATATATTACAGGCTTTATCTTACGTCAGGCCATT 1116
Db 1983 AspMetValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValGlyValIle 1902
Qy 1115 GGAACCTCAATGTGGAGACTGTTTGTAAATCAGACAT-----GGAGAGGCTGCA--- 1068
Db 1903 CysAlaAlaIleLeuArg-----ArgHisValGlyProGlyGluGlyAlaVal 1918
Qy 1067 -----CGTCTAAGCGGAGACAGCTGCTTCGTTGGGAATCATCAGCTCCCTCCGCT 1014
Db 1919 GlnTrpMetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHis 1938
Qy 1013 CAC-----GCCGCTCTTCCCTCCCGCTGTTTC 984
Db 1939 TyrValProGluSerAspAlaAlaArgValThrAlaLeu--LeuSerSerLeuThrV 1958
Qy 983 ACAGCTGTTCCAGATTTGTCAGCAAGAAATAATGAA-----TGCA 939
Db 1958 alThrSerLeuLeuArgLeuHisGlnTrpIleAsnGluAspTyrProSerProCys 1978
Qy 938 TACAGGACTTT-----928
Db 1978 erAspAspTrpLeuArgIleIleTrpAspTrpValCysSerValLeuThrAspPheLys 1998
Qy 927 --TGCTA-----GTAGACTGTCTG- 910
Db 1998 hrTrpLeuSerAlaLysIleMetProAlaLeuProGlyLeuProPheIleSerCysGlnL 2018
Qy 909 -----GGTATTGTGAGCATGCGAGTGTGTTGTTTCTG 879
Db 2018 ysGlyTyrLysGlyValTrpArgGlyAspGlyValMetSerThrArgCysProCysGlyA 2038
Qy 878 TTATTATAATAAAGTCAAAACGTGAGTCAAC-----847
Db 2038 laThrIleThr-----GlyHisValLysAsnGlySerMetArgLeuAlaG 2053
Qy 846 -----ACTTTCCCATCTTCTGTTTATACACAA 819
Db 2053 lyProArgThrCysAlaAsnMetTrpHisGlyThrPheProIleAsnGluTyrThrThrG 2073
Qy 818 GTCAAACTTCTCATTCATGAGACGGTGAAGATGCTCTCCACAGCCCAACTCTTCCCG 759
Db 2073 ly-----ProSerThr-ProCysProSerPro 2081
Qy 758 GCCCAGCGGAGGGCTC-----ACGGTTGACAGCTGTCTCCACAGCTGGTGGCAGTCTC 702
Db 2082 AsnTyrThrArgAlaLeuTrpArgValAlaAla-----AsnSerTyrValGluValArg 2099
Qy 701 CGTCAGGCTTCCAGCTGGGCTTGTCCACAGGACACCCCGAGGAATCTGCGCCAGCTCTC 642
Db 2100 ArgValGly---AspPheHisTyrIleThrGlyAlaThrGluAspGlu-----Leu 2115
```

```
Qy 641 CACCATCGTCACAGGTCCTCCGATCGATCTCTTATCTTGAGAAAAGCAGCGTTCAGTC 582
Db 2116 LysCysProCysGlnValProAlaAlaGluPheThrGluValAspGlyValArgLeu 2135
Qy 581 CAT----- 579
Db 2136 HisArgTyrAlaProProCysLysProLeuLeuArgAspGluIleThrPheMetValGly 2155
Qy 578 -----GCGGTG 573
Db 2156 LeuAsnSerTyrLeuIleGlySerGlnLeuProCysGluProGluProAspValSerVal 2175
Qy 572 CTCCAGAACTCCG-----CACGTGCTCAAAACACGAGCGGTAGCCAGCTT 525
Db 2176 LeuThrSerMetLeuArgAspProSerHisIleThrAlaGluThrAlaAlaArgArgLeu 2195
Qy 524 ATCATTCATAAACCTCCGGCAGAAATCTTGAAGAGTCCGTCGGTAGCTCATGTGTCGCAG 465
Db 2196 AlaArgGlySerProProSerGluAlaSerSerSerAlaSerGlnLeuSerAlaPro--- 2214
Qy 464 AGACGGTGGAACTGATAAAGACACACCATCATCTTGGGGTTCGAGCCATATAGAT 405
Db 2215 ---SerLeuLysAlaThrCysGlnThrHisArgProHisProAspAlaGluLeuValAsp 2233
Qy 404 -----GACCTTGGAGTCTCCATTGTGGAGGTTCAGAGGGCAGAAA 366
Db 2234 AlaAsnLeuLeuTyrArgGlnGluMetGlySerAsnIleThrArgValGluSerGluThr 2253
Qy 365 GCGGTAGGCGAGGTGGTCTTGTATGAGCGGGGAGAGGTTCAGTTCTTGATGTGTCAG 306
Db 2254 LysVal-----ValIleLeuAspSer-----PheGluProLeuArg 2265
Qy 305 GCCGGGTGGGTACTCCAGACGGGAGGTGTCTGTGCA-----TGTTCAATCAAGC 253
Db 2266 AlaGlu-----ThrAspAspAlaGluLeuSerValAlaAlaGluCysPhe----- 2280
Qy 252 CGATCTCATCGGGTCAGCGCCTGGCTCACAAGTAGACACCTCTCGCAGCAAGCTGG 193
Db 2281 -----LysLysPro-----ProLysTyrProProAlaLeuProIleTyr 2293
Qy 192 TGCGGAGCTTGGGTAGTGACGATCCACACGTCTGGCGCCGACCGGGAAGTTGGCGA 133
Db 2294 AlaArgProAspTyrAsnProProLeuLeuAspArgTyrLysAlaProAsp-----Tyr 2311
Qy 132 TCTCTCATCTTCCCGCGGAGAGGGCGCAGCCGACGCCATCGAAGTACT 73
Db 2312 ValProProThrValHisGly----- 2318
Qy 72 TGCTCTGMACTCCCGGGGTGCTGGGGGTCTCGGCTCGCTCTCGCCATGCGCGC 13
Db 2319 CysAlaLeuProProArgGlyAla-----ProProValProProProArgArg 2334
Qy 12 CGTCGC 7
Db 2335 LysArg 2336

RESULT 27
US-10-196-749-238
; Sequence 238, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
```

```
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 238
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-238

Alignment Scores:
Pred. No.: 0.461 Length: 325
Score: 96.00 Matches: 70
Percent Similarity: 29.5% Conservative: 20
Best Local Similarity: 23.0% Mismatches: 104
Query Match: 2.2% Indels: 111
DB: 6 Gaps: 15

US-10-768-158-1 (1-2419) x US-10-196-749-238 (1-325)
Qy 850 TCACACTTTCCTCCATCTTCTTTATATACACCAAGTCAAACTTCTCATTCTATGAGACGGTG 791
Db 58 SerGluHisProGlnProAlaLeuAspProArgSerAsnAspLeuAlaArgValProLeu 77
Qy 790 AAGATGTCCTTCCACAGCCCAACTCTTCCCGGCCACGCGGAGGGCTCAGCGTTGCAG 731
Db 78 LysLeuSerVal---ProProSerAspGlyPheProProAlaGlyGlySerAlaValGln 96
Qy 730 CACTGGTCCACCAAGCTGG-----TGG 710
Db 97 ArgTyrProProSerTyrGlyLeuProAlaMetAspSerTyrProProGluAspProTyr 116
Qy 709 CAGTGTCCGTCAGGGCTTCAGCTGGCGCTTGTACAGCAGCACCCCGCAAT-----656
Db 117 GlnMetMetAlaAlaAlaAlaGluAspArgLeuGlyGluAlaLeuProGluGluLeuSer 136
Qy 655 ---CTGCCAGCTGCTCCACCATCGTCACAGGTCCCGATGTCATGTCTTCTCATCTGAGA 599
Db 137 TyrLeuSerSerAlaAlaLeu-AlaProGly-----147
Qy 598 AAAAGCAGCTTCGAGTCCATGCGGTGCTCCAGAACTCTCCAGAACTCTCCAGTCTCAACACGAG 539
Db 148 -----SerGlyProLeuProGlyGluSerSerProAspAlaThrGlyLeuSerPr 164
Qy 538 CCGTAGCCACGCTTATCATTCATAAACCCTCCGGCAGAAATTTGAAAGGTGCCTCGGTAG 479
Db 164 oGluAlaSerLeuLeuHisGlnAspSerGluSerArgArgLeuProArg-----180
Qy 478 CTCATGTCGCGCAGAGCGGTGGAATGTGATAAAGACACACCATGATCTTGGGGTTG 419
Db 181 -SerAsnSerLeuGlyAlaGlyGlyLysIleLeuSerGlnArgPro-----195
```

Qy	418	CGAGCCATATAGATGACCTTGGAGTCTC-----CATTTGGCAGG	380
Db	196		
Qy	379	TCAGAGGCAGAAACGGTAGGGCAGGTGGCTCTTTGATGAGCGGGAGAGTCAAGTTC	320
Db	210	yThrLeuAsnProSerValSerTrpGlyGly	220
Qy	319	TTGATGATGTCACGCCCGGTGCTGGTACTCCAGGACCGGAGCTGCTCGTGATGTC	260
Db	221	-----GlyGlyProGlyThrGly-----	226
Qy	259	ATCAAGCCGATCTCATCGGGGTACGCGCCTGGCTCACCAAGTAGACCACTCTCGCAGC	200
Db	227	-----TrpGlyThrArgPro-----MetProHisProGluG1	237
Qy	199	AAGCTGG-----TGC CGGACTTGGGTAGGTGACGATCCACACG	161
Db	237	yIleTrpGlyIleAsnAsnGlnProProGlyThrSerTrpGly-----AsnIleAs	254
Qy	160	TCGCTGGCGCCGACCGGGAAGTTGG-----CGATCTCTCCATCTCCCGCGG	113
Db	254	nArgTy+ProGly---GlySerTrpGlyAsnIleAsnArgTy+ProGlyGlySerTrpG1	273
Qy	112	CAGAAAGGCGCGACCGCCGACGCCATCGAAGTACTTGCTCTCGCAACTCCCCCGG	53
Db	273	y-----AsnIleAsnArgTy+ProGlyG1	281
Qy	52	GTGCTGGGGGTCT	40
Db	281	ySerTrpGlyvAsn	285

RESULT 28

```

US-11-101-316-64
; Sequence 64, Application US/1101316
; Publication No. US20060099657A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
; TITLE OF INVENTION: UNDEREXPRESSED IN MELANOMA
; FILE REFERENCE: P3230R1C17C1
; CURRENT APPLICATION NUMBER: US/11/101.316
; CURRENT FILING DATE: 2005-04-06
; PRIOR APPLICATION NUMBER: 10/063526
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 03/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 64
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-101-316-64

```

Alignment Scores:

Pred. No.:	0.461	Length:	325
Score:	96.00	Matches:	70
Percent Similarity:	29.5%	Conservative:	20
Best Local Similarity:	23.0%	Mismatches:	104
Query Match:	2.2%	Indels:	111

DB: 7 Gaps: 15

US-10-768-158-1 (1-2419) x US-11-101-316-64 (1-325)

Qy	850	TCACACTTTCCCATCTTCTGTTTTATACACCAGTCAAACTTCTCATTCTCATGGAGACGGTG	791
Db	58	SerGluHisProGlnProAlaLeuAspProArgSerAsnAspLeuAlaArgValProLeu	77
Qy	790	AAGATGTCTTCCACAGCCCAACTCTTCCC CGGCCACCGGCAGGCGCTCACGCTTGACG	731
Db	78	LysLeuSerVal--ProProSerAspGlyPheProProAlaGlyGlySerAlaValGln	96
Qy	730	CAC TGGTCCACCAGCTGG-----TGG	710
Db	97	ArgTrpProProSerTrpGlyLeuProAlaMetAspSerTrpProProGluAspProTrp	116
Qy	709	CAGTGTCCGTCAGGGCTTCAGCTGGGCCCTTGTACACAGGACACCCCCAGGAAT----	656
Db	117	GlnMetAlaAlaAlaGluAspArgLeuGlyGluAlaLeuProGluLuuSer	136
Qy	655	--CTGGCCAGCTCTCCACCACCTCGTACCAGGTCGCCGATGCATGCTCTCATCATTTAGA	599
Db	137	TyrLeuSerSerAlaAlaLeu-AlaProGly-----	147
Qy	598	AAAAGCAGTTGAGTGCATCCGCTGCCAGAACCTCTGCACGTCTCAAACACAGGAG	539
Db	148	-----SerGlyProLeuProGlyGluSerSerProAspAlaThrGlyLeuSerPr	164
Qy	538	CCGTAGCCCCAGCTTATCATTTATAA CCTCCGGCAGATTCTTGAAGGTGCCTCGGTAG	479
Db	164	oGluAlaSerLeuLeuHisGlnAspSerGluSerArgArgLeuProArg-----	180
Qy	478	CTCATGTGTCCGACAGCGGTGGAACCTGATAATAAGACACACCACCATCATCTTGGGTTG	419
Db	181	-SerAsnSerLeuGlyAlaGlyGlyLysIleLeuSerGlnArgPro-----	195
Qy	418	CGAGCCATATAGTAGACTTGGAGTCTC-----CATTTGGAGG	380
Db	196	-----ProTrpSerLeuIleHisArgValLeuProAspHisProTrpGl	210
Qy	379	TCAGAGGGCAGAAAGCGTAGGGCAGGTGGCTCTTGATGAGGCGGGAGAGGTCA GTTC	320
Db	210	yThrLeuAsnProSerValserTrpGlyGly-----	220
Qy	319	TTGATGATGTCACAGCCCGGTGTGGTGTACTCCAGAGCCGGAGCTGCTGCTCGATGTC	260
Db	221	-----GlyGlyProGlyThrGly-----	226
Qy	259	ATCAAGCCGATCTCATCGGGTCAGCGCTTCAGGTCTCAACAAGTAGACCCTCTCAGC	200
Db	227	-----TrpGlyThrArgPro-----MetProHisProGluGl	237
Qy	199	AAGCTGG-----TGCCGGACTTGGGGTAGGTACGATCCACAGC	161
Db	237	yIleTrpGlyIleAsnAsnGlnProProGlyThrSerTrpGly-----AsnIleAs	254
Qy	160	TGCGTGGCGCCACCGGGAAGTTCG-----CGATCTCTCCATCTTCCCGGG	113
Db	254	naGtyrProGly---GlyserTrpGlyAsnIleAsnArgTyrrProGlyGlySerTrpGl	273
Qy	112	CAGAAGGGCGGCAGCCGACGCCATGGAACCTCGAAGTACTTGCTCTCGAACTCCCCCGGG	53
Db	273	y-----AsnIleAsnArgTyrrProGlyGl	281
Qy	52	GTGCTGGGGTCT	40
Db	281	ySerTrpGlyAsn	285

RESULT 29  
 US-10-199-229-13  
 ; Sequence 13, Application US/10199229  
 ; Publication No. US2006099701A1  
 ; GENERAL INFORMATION:

```
; APPLICANT: She, Jin-Xiong
; APPLICANT: Kumar, Pradeep
; APPLICANT: Wang, Cong-Yi
; TITLE OF INVENTION: ACTIVATORS OF CYCLIN-DEPENDENT KINASES
; FILE REFERENCE: 5853-210
; CURRENT APPLICATION NUMBER: US/10/199,229
; CURRENT FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-199-229-13

Alignment Scores:
Pred. No.: 0.85 Length: 951
Score: 95.00 Matches: 69
Percent Similarity: 34.3% Conservative: 23
Best Local Similarity: 25.7% Mismatches: 95
Query Match: 2.2% Indels: 81
DB: 6 Gaps: 13

US-10-768-158-1 (1-2419) x US-10-199-229-13 (1-951)
Qy 783 CTTCCACAGCCCACTCTCCCGGCCACGGGCGGCTCAGCGTTGTCAGCACTGGT 724
Db 81 ProValProAlaAlaProValProAlaProSerLeuAlaProGlyGluAenGlyThrGly 100
Qy 723 -----CCACCAGCTGGTGGCAGTGCTCCG 700
Db 101 AspTrpAlaProArgLeuValPheIleGluProProGlyAlaGlyGlyAlaAlaPro 120
Qy 699 TCAGGGCTTCAGCTGGGCGCTTGTACAGGACACCCCGGAGAACTGGCCAGCTGCTCCA 640
Db 121 SerAlaValProThrArgPro----- 127
Qy 639 CCATCGTCACAGGTCGCGATGCTTTCATCTTGAGAAAAGCAGCTTCGAGTCCA 580
Db 128 -----ProGlyProGlnArgCys-----ArgGluGlnSerAsp-Tr 139
Qy 579 TGGGTGTCTCCAGAACTCTCGACGCTGCTCAACACGAGCGGTAGCC---CAGCTTAT 523
Db 139 pAlaSerAspValGluValLeuGlyProLeuArgProGlyGlyValAlaGlySerAlaLe 159
Qy 522 CATTCAATAACTCCGGCAGAACTTTGAAAGGTGCTCGGTAGCTCATGTTCCGACAG 463
Db 159 uValGlnValArgValArgGluLeuArgGlyGlyGluAlaGluArgGlyGlyAlaGlyG 179
Qy 462 AGCGGTGNACTGATAATA-----AGACACACCCAG-- 432
Db 179 yGlyGlyLeuPheSerLeuCysAlaTrpAspGlyArgAlaTrpHisHisGlyAl 199
Qy 431 -----ATCCTTGGGGTTCGCGACCATATAGATACCTTGGAGTCTCCATTGT 385
Db 199 aAlaGlyGlyPheLeuLeuArgValArgProArgLeuTyrglyProGlyGlyAspLeuLe 219
Qy 384 GGAGGTCAGAGGGCAGAAAGCGGTAGGCGAG-----GTGGCTCTTGATAGGCGGG 334
Db 219 uProAlaTrpLeuArgAlaLeuGlyAlaLeuLeuLeuLeuAlaLeuSerAlaLeuPh 239
Qy 333 GAGAGGT-----CAGTCTTGATCATGTCAGGCGGCTCATCGGGGTGAGCGGCT 289
Db 239 eSerGlyLeuArgLeuSerLeuLeuSerLeuAspProValGluLeuArg-----ValLe 257
Qy 288 CCAGACCGGAGCTGCTCGTGCATGTTTCATCAAGCCGATCTCATCGGGGTGAGCGGCT 229
Db 257 uArgHisSerGlySerAlaAlaGluGlnGluGlnAlaArg-----ArgValGlnAlaVa 275
Qy 228 GGCTCACCAGAGTAGACAC-----CTCCTGCAGCAAGCTGGTGGCGGACT 184
Db 275 lArgGlyArgGlyThrHisLeuLeuLeuLeuLeuLeuGlyGlnAlaGlyAlaAsnAl 295
```

```
183 T-----GGGCTAGGTGAC-----GA 169
295 aAlaLeuAlaGlyTrpLeuTyraSerLeuProGlyValGlyAspProGlyGluAs 315
Qy 168 TCCACACGTCGCTGGCGGCAC 147
Db 315 pSerGlyGluAlaGlyValHis 322

RESULT 30
US-10-505-928-150
; Sequence 150, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 150
; LENGTH: 5738
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-505-928-150

Alignment Scores:
Pred. No.: 1.72 Length: 5738
Score: 95.00 Matches: 152
Percent Similarity: 29.3% Conservative: 76
Best Local Similarity: 19.6% Mismatches: 293
Query Match: 2.2% Indels: 257
DB: 6 Gaps: 31

US-10-768-158-1 (1-2419) x US-10-505-928-150 (1-5738)
Qy 94 TCGCGCTCGCCCTTCTGCGCGGGAAGATGGAGGAGATCGCAACTTC----- 144
Db 4211 CysAspCysThrAspAlaGlnGlyHisSerTrp-----AlaProGlySerGlnHisGln 4228
Qy 145 -----CGGTGGCGCCAGCAGCTGTGGATCGTCACTACCTACC 180
Db 4229 AspAlaCysAsnAenCysSerCysGlnAlaGlyGlnLeuSerCysThrAlaGlnProCys 4248
Qy 181 CCAAGTCGGCACCAGCTTGTGACGAGGTGTCTACTTGTGTGAGCAGGCGGCTGACC 240
Db 4249 ProProThrHisCysAlaTrpSerHisTrpSerAlaTrp-----Ser 4263
Qy 241 CCGATGAGATCGGCTTGTATGAACATCGACGAGCAGCTCCCGGTCTCGAGTACCACAGC 300
Db 4264 ProCys-----SerHisSerCysGlyProArgGlyGlnGlnSerArgPhe 4278
Qy 301 CGGCGC-----TGGACATCATCAAGAACTGACCTCTCCCTC-----GCCTCATCA 345
Db 4279 ArgCysGlyProGlyLeuAlaSerArgSerCysProCysLeuMetAlaLysAla 4298
Qy 346 AGAGCCACCTGCCCTACCGTTTC----- 369
Db 4299 AspProThrCysAenSerThrPheLeuHisLeuAspThrGlnGlyCysTySerGlyPro 4318
Qy 370 TGCCTCTGACCTCCACAATGGAGACTCCAAGGTCACTATATGGCTCGCAACCCCAAGG 429
Db 4319 CysProAspSerCys-GlnTrpSerLeuTrpGlyProTrpSerProCysGlnValProCy 4338
Qy 430 ATCTGGTGGTCTCTATTATC-----AGTTCACCGCTCTCTCGGACCATGAGCTTAC 482
Db 4338 sSerGlyGlyPheArgLeuArgTrpArgGluAlaGluAlaLeuCysGlyGlyGlyCysAr 4358
Qy 483 CGAGGCACTTTTCAAGAAATTCGCGGAGGTTTATGAATGATAGCTGGGCTCGGCTCC 542
Db 4358 gGlu-----ProTrpAlaGlnAspAr 4365
```

QY 543 TGGTTTG-----AGCAGTCGACGAGTCTCTGGAGCACCGCATGGAC 584  
Db |||||  
4365 glysLeuGlnArgArgAlaLeuProSerThrCysValAsnGluSerLeuValCysProHi 4385  
QY 585 TCGAACGTGCTTTTCTCAAGTATGAACACATGCATCGGACCTGGTACGATGGAG 644  
Db |||||  
4385 sGlnGluCys-----ProValLeuGlyProTrpSerAlaTrpSerSe 4399  
QY 645 CAGCTGGCCAGATTCTCTGGGGTGCTCTGCACAGG-----CC 683  
Db |||||  
4399 rCysSerAlaProCysGlyGlyThrMetClnuArgHisargThrCysGluGlyGlyPr 4419  
QY 684 CAGCTGGAGCCCTGACGAGCAGCTGCACAGCTGGTGGACAGTCTGCAACCTGAG 743  
Db |||||  
4419 oGlyValAlaProCysGlnAlaGlnAspThrGlu---GlnArgGlnGluCysAsnLeuG1 4438  
QY 744 GCCCTGCCCGTGGGCC-----GGGGAAGATGGCTGTGGAGGACATCTTACCG-- 795  
Db |||||  
4438 nProCysProGluCysProProGlyGlnValLeuSerAlaCysAlaThrSerCysProCy 4458  
QY 795 ----- 795  
Db  
4458 sLeuCysTrpHisLeuGlnProGlyAlaAlaCysValGlnGluProCysGlnProGlyCy 4478  
QY 796 -----TCTCCATGAATGAGAAGTTTGACTTGTGTATAAACAGAGATGGGAAAGTGT 848  
Db |||||  
4478 sGlyCysProGlyGlnHisSerLeuProTrpGlyLeuThrLeuGluGluG1 4498  
QY 849 GACCTCAGTTTGACT-----TTTATTATTAATAACAGAAACACACCTGCATGCTCA 902  
Db |||||  
4498 nAlaGlnGluLeuProProGlyThrVal-LeuThrArgAsnCysThrArgCysValCysH 4518  
QY 903 CAATACCCAGACAGCTACTAGCAAAAGTCTGTATGCATTTCATTTATCTTCTGCTGGA 962  
Db |||||  
4518 isGlyGlyAlaPheSerCysSerLeuValAspCysGlnGlyValPro-ProGly 4537  
QY 963 CAAACTCTGGAAGCA----- 977  
Db |||||  
4538 GluThrTrpGlnGlnValAlaProGlyGluLeuGlyLeuCysGluGlnThrCysLeuGlu 4557  
QY 978 -----GGCTGTGAACAGCGGGGAAGGAGCGCGTGGAGGGAGGTGTGATG 1031  
Db |||||  
4558 MetAsnAlaThrLysThrGlnSerAsnCysSerSerAlaArgAlaSerGlyCysValCys 4577  
QY 1032 ATTCCTCAACCGAAGCAGCTGTCTGCCCTTTAGAACGTGCAGCTCTCCATGCTGTGATTAC 1091  
Db |||||  
4578 GlnProGlyHis-----PheArgSer-GlnAlaGlyProCysValProG1 4592  
QY 1092 AAACAGTCTCCACATTGCAGTTCCAAATGGCCGTGACCGTAAGGATAAGCCTGTAATATA 1151  
Db |||||  
4592 uAspHisCysGlu-----CysTrpHisLeuGlyArgProHisLeuValArg 4607  
QY 1152 TGCACCTAGATGTCTGCTTTTCAACCCCGTATTTATGATTTATATAGACTTTTCACT 1211  
Db |||||  
4607 gHisArgThrProSerAlaThrThrHisProPheLeu-----ThrProSerLe 4623  
QY 1212 GGAATCTACATAAATGTCTAGTAAACCAACAAATAAAGTTTCATTTCCAGGGGAATCAGGAG 1271  
Db |||||  
4623 uProlleCysLeuGlyArgGlyProGluProTrpLeuGlySerValGlnGlyGlnArgSe 4643  
QY 1272 C-----GAGCCACACCCGAATGGTGAAGAGATCTCAGGGTTAACTCTTTATTTTGTAG 1325  
Db |||||  
4643 rLeuLeuGluHisPhePro---GlyLysLys----- 4652  
QY 1326 TTTTATTATCTAAGGCACAGCCATCTGTTCTCACTGTTCTGAGATAGTGGTGAGAAC 1385  
Db ----- 4652  
QY 1386 AGAGATGAGTTGGTCTGTTGGGGGAATCTGGACACTTCTTTTATTCTGACGGAGTTCA 1445  
Db |||||  
4653 -----GlnMetGlyTyrMetGly----- 4658

QY 1446 CTTCTTCAGAACCTTCTCTGAAATGAGCAGAAATGTTTCACTAGGTCTTTCAGAAATGGACGT 1505  
Db |||||  
4659 -----GluMetGluVa 4662  
QY 1506 CTTCTGCGCAGAGACTTCCAGCGGCGCTCCAAAGCCCAATGCAGAGGAGCGCGCGA 1565  
Db |||||  
4662 lGlnGlyProThrArgGluSerGlyGlnSerLeuProProGln----- 4676  
QY 1566 GCATGTGCTGAGGAAGTCTGCTGCTGGTGGAGCTGGCAGGTGGAGTCTTAATGCACTCAGG 1625  
Db |||||  
4677 -----LysLysAlaTyrLeuSerHisLeuSerThrGlySerGlyHisIleGluG1 4693  
QY 1626 AGCATTTTCATGCTGGTGGAGAGTGGCGGCCCAAGACCCAGGTCCTCGGAATT 1685  
Db |||||  
4693 yAspTrpAlaGlyArgAsnArgLysLeuLysLeuLysProArgSerIleGlnLysSerTrpPh 4713  
QY 1686 T-----GAGCTGAATTCACACGCTTACTTTGTTTCTCTGAA 1721  
Db |||||  
4713 eValGlnPheProTrpLeuIleMetAsnGluGluGlnThrAlaLeuPheCys----- 4730  
QY 1722 GTGATAGCTACTAATCTGCAAGCAGATGCTTTAATAGTAAATTTCTAAATATCCCGGG 1781  
Db |||||  
4731 -----SerAlaCysArgGluTyrProSerIleArgAspLysArgSerArgLe 4746  
QY 1782 TCCT-----TATCATTCAGTTTGTTC 1802  
Db |||||  
4746 uIleGluGlyTyrThrGlyProPheLysValGluThrLeuLysTyrHisAlaLysSerLy 4766  
QY 1803 TGTCCACTGAGGCGCTCAGCGTGGGAGGACCAATTTTGGAGTGTGACCTGTTTCACT 1862  
Db |||||  
4766 sAlaHisMet-----PheCysValAsnAlaLeuAlaLaAr 4778  
QY 1863 CGATCAGTTTGGCAGCGCGCTGCTGTCTGCCACCTCATCCCTCGTGTATCTGAG 1922  
Db |||||  
4778 gAspProIleTrpAlaalaArgPheArg-SerIleArgAspProProGlyAspValLeuA 4798  
QY 1923 GGAGTAAAGTGAGGTCTTTATGCTTCACTGCTCACTGCTCAATTTTCTCACC----- 1970  
Db |||||  
4798 laSerProGluProPheThrAlaAsp-CysProIlePheTyrProProGlyProLeu 4817  
QY 1971 -----ACATTCTGCTGAAGCGATGAGAGTGGGGCGGCGCAGTAGCCAGCCACACCC 2018  
Db |||||  
4818 GlyGlyPheAspSerMetAlaGluLeuLeuProSerSerArgAlaGluLeuGluAspPro 4837  
QY 2019 CGTGGGACCGGGTCTCTCTCATTTATGTTGGTGGAA 2057  
Db |||||  
4838 GlyGlyAsp---GlyAlaIleProAlaMetTyrLeuAsp 4849

## RESULT 31

US-10-196-749-284  
; Sequence 284, Application US/10196749  
; Publication No. US2006094864A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Goddard, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3430R1C340  
; CURRENT APPLICATION NUMBER: US/10/196,749  
; PRIOR FILING DATE: 2002-07-16  
; PRIOR APPLICATION NUMBER: 10/052586  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18





```

Db      382 rSerArgGlyProSerAspThrSerSerArgPheSerLysGluGlnArgGlyValAlaG1 402
Qy      674 -----TGCAAGGCCAGGTGGAGCCCTGACGG 702
Db      402 nProLeuAlaGlnLeuArgSerCyeProGlnGluGluGlyProArgGlyArgGlyLeuAl 422
Qy      703 AGCACTGCCA-----CCAGCTGGTGGACCACTGCTGCACACGCTGAG----- 743
Db      422 aAlaArgProLeuGluAenArgAlaGlyGlyProVal-AlaArgSerGluGluProGlyA 442
Qy      744 --GCCTCCCGTGGCGCGGGAAGATT-----GGGCTGTGGGAAGACATCTTCA 792
Db      442 laProLeuProValAlaValGlyThrAlaGluProGlyGlySerMetLysThrThrPheT 462
Qy      793 CCGTCTCCATGAATGAG 809
Db      462 hrIleGluIleLysAsp 467

RESULT 33
US-10-525-621-1
; Sequence 1, Application US/10525621
; Publication No. US20060100418A1
; GENERAL INFORMATION:
; APPLICANT: Kiyosue, Yuko
; APPLICANT: Sasaki, Hiroyuki
; APPLICANT: Tsukita, Shoichiro
; APPLICANT: Eisai Co., Ltd.
; TITLE OF INVENTION: CULTURED XENOPUS LAEVIS CELL LINES
; TITLE OF INVENTION: EXPRESSING MUTANT ADENOMATOUS POLYPOSIS COLI GENE
; FILE REFERENCE: 082368-002400US
; CURRENT APPLICATION NUMBER: US/10/525,621
; PRIOR FILING DATE: 2005-02-22
; PRIOR APPLICATION NUMBER: PCT/JP03/10434
; PRIOR FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: JP 2002-241487
; PRIOR FILING DATE: 2002-08-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2829
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-10-525-621-1

Alignment Scores:
Pred. No.: 1.73 Length: 2829
Score: 93.50 Matches: 82
Percent Similarity: 32.0% Conservative: 26
Best Local Similarity: 24.3% Mismatches: 140
Query Match: 2.1% Indels: 89
DB: 6 Gaps: 15

US-10-768-158-1 (1-2419) x US-10-525-621-1 (1-2829)
Qy      840 CCATCTTCTGTTTATACCAAGTCAAACTTCTCATTCATGAGACGGTGAAGATGCTCT 781
Db      2286 ProSerSer-----ThrProGlyCysSerLysGlyAsnSerArgSerGlySerArg 2303
Qy      780 TCCACAGCCCAACTCTTCCCGGCCCCACGG-----GCAGGGCCCTCAG 739
Db      2304 AspSerAlaSerArgProSerProGlnProLeuSerArgProLeuGlnSerProGly 2323
Qy      738 CGTTGCACACTGGTCCACCAAGCTGGTGGCAGTGTCTCG-----TCAGGGGCTT 691
Db      2324 ArgAsnSerIleSerProGlyLysAsnGlyIleSerProAsnLysPheSerGlnLeu 2343
Qy      690 CCA-----GCTGGCCCTTGTGCAC 673
Db      2344 ProArgThrThrSerProSerThrAlaSerThrLysSerSerGlySerGlyArgMetSer 2363
Qy      672 AGGACACCCCGAGGATCTGCCAGCTCTCCA----- 640

```

```

Db      2364 TyrThrSerProGlyArgGlnLeuSerGlnProAsnLeuSerLysGlnSerGlyLeuPro 2383
Qy      639 -----CCATCTCTCACAGGTCCCGATGCCATGCTCTTCAT----- 607
Db      2384 LysThrHisSerSerIleProArgSerGluSerAlaSerLysSerLeuAsnGlnAsnVal 2403
Qy      606 -----ACTTGAGAAAAGACAGCTTCGAGTCCATCGCGGTGCTCCAGCAACTCCTGCA 556
Db      2404 AsnThrGlySerAsnLysLysValGluLeuSerArgMetSerSerThrLysSerSerGly 2423
Qy      555 CGTCTCAAAACAGGAGCGGTAGCCACGCT-----TATCATTCATAA 514
Db      2424 SerGluSerAspArgSerGluArgProAlaLeuValArgGlnSerThrPheIleLysGlu 2443
Qy      513 ACCTCCGGCAGAAATTCGTGAAGTCTCGGTAGTCATGGTCCGAGAGAGCGGTGGA 454
Db      2444 AlaProSerProThrLeuArgArgLysGluLeuSerAlaSerPheGluSerLeuSer 2463
Qy      453 ACTGATAATAAG-----ACACCACAGATCCTTGGGGTTCCGAGCCATATAGATGACCT 400
Db      2464 SerSerSerArgAlaAspSerProProArgSerGlnThrGlnThrProAlaLeuSerPro 2483
Qy      399 TGGAGTCTCCATTTGTGGAGGTACAGGGCGCAAAAGCGGTAGGCGAGGTGGCTCTTGATGA 340
Db      2484 ---SerLeuProAspMetAlaLeuSerThrHisSerIleGlnAla----- 2497
Qy      339 GCGCGGAGAGGTGAGTTCCTTGATGATGTCCAGGCCGCGCTGTGGGTACTCTCCAGGACCG 280
Db      2498 GlyGlyTrpArgLysMetPro-----ProAsnLeuAsnProAlaAlaGluHisGly 2514
Qy      279 GGACCTCTCTCGTCAATGTTTCATCAAGCGGATCTCATCGGGTACGATCCACACGT 160
Db      2515 AspSerArgArgArgHisAspIleSerArgSerHis-----SerGluSerPro 2530
Qy      219 AGTAGACCACTCTCTCGCAGCAAGCTGTCGCGGACTTGGGTAGGTGACGATCCACACGT 160
Db      2531 SerArgLeuProlleThrArgSer-----GlyThrTrpLysArgGluHisSerLysHis 2548
Qy      159 CGCTGGGCGCCACCG-----GGAAGTTGGCGATCTCTCCATCTTCCCGCGGAGAGG 106
Db      2549 SerSerSerLeuProArgValSerThrTrpArgArgThrGlySerSerSerIleLeu 2568
Qy      105 GCGGACGCGCACGCCAT----- 88
Db      2569 SerAlaSerSerGluSerSerGluLysAlaLysSerGluAspGluLysGlnValCys 2589
Qy      87 -----GGAAGTCCGAAGTACTTGTCTCGAACTCCCGCGGTGCTGG 46
Db      2589 SerPheProGlyProArgSerGluCysSerSerSerAlaLysGlyThrTrp 2605

RESULT 34
US-10-505-928-851
; Sequence 851, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 851
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-851

Alignment Scores:
Pred. No.: 0.847 Length: 353
Score: 93.00 Matches: 81

```

```
Percent Similarity: 29.7%      Conservative: 24
Best Local Similarity: 22.9%    Mismatches: 113
Query Match: 6                 Indels: 135
DB:                             Gaps: 18

US-10-768-158-1 (1-2419) x US-10-505-928-851 (1-353)

Qy 1 CGGAGCGGCGGCGGCGGATGG-----CGGAGCGGAGCGCG 39
Dy 15 AlaArgProAnaAlaGlyAlaTrpSerIleGluAlaGlyProSerArgAspSerArgAla 34
Qy 40 AGACCCCGCAGC-----CCCGCGGCGAGTTCGAGCAGCAAGTACTTCGAGT 84
Dy 35 TrpProAspAlaGlyAsnHisThrLeuAlaGlnThrAlaSerProAspGlyThrGluPro 54
Qy 85 TCCATGGCGTGGCTGCGCGCTTCTGCCCGGGAAGATGGAGAGATCCCAACT--- 141
Dy 55 GlyHisSerProGlyCysLeuAsnSerAlaGly-----ValSerThrProThrSer 71
Qy 142 ---TCCCGGTGCGGCCA----- 156
Dy 72 GlySerMetCysProProGlnAlaGlnAlaGluValGlyProThrMetThrGluLeuAla 91
Qy 157 ---GCGAGCTGTGATCGTCACTACCCAGTCCGCGCACCAGCTTGCTGCAGGAGGTG 213
Dy 92 GluMetValCysAlaProSerProAlaProAlaProProLys-ProAla----- 108
Qy 214 TCTACTTGGTGGAGCGCGCTGACCCGATGATCGGCTTGATGAACATCGACGAGC 273
Dy 109 -----SerProGlyPro--ProGlnValGluGluValGlyHisArgGlyGlySe 124
Qy 274 AGCTTCCC-----GCTCTGGAGTACCCAGCGCGGCTGGACATCATCAAGNAC 324
Dy 124 rSerProProArgLeuProProGlyValProValIleSerLeuGlyHisSerArgProPr 144
Qy 325 TGACCTTCCCGCTCATCAAGAG-----CCACCTGCCCTACCGCTTTC 369
Dy 144 oGlyValAlaMetProThrThrGluLeuGlyThrLeuArgProProLeuLeuGlnLeuSe 164
Qy 370 TGCCCTC-----TGACCTCCCAATGAGACTCCAAGTCACTATAT----- 412
Dy 164 rThrLeuGlyThrAlaProProThrLeuAlaLeuHisTyHisProHisProPheLeuAs 184
Qy 412 ----- 412
Dy 184 nSerValTyrlleGlyProAlaGlyProPheSerIlePheProSerSerArgLeuLysAr 204
Qy 413 -----GGCTCGCAACCC----- 424
Dy 204 gArgProSerHisCysGluLeuAspLeuAlaGluGlyHisGlnProGlnLysValAlaAr 224
Qy 425 -----CAAGGATCTGGTGGTGTCTTATT 447
Dy 224 gArgValPheThrAsnSerArgGluArgTrpArgGlnGlnAsnValAsnGlyAlaPheAl 244
Qy 448 ATCAGTCCAGCTCTCTGGCGACCATGAGTACCGAGGACCTTTCAAGAATCTCTGCC 507
Dy 244 aGluLeuArgLysLeuLeuProThrHisProProAspArgLysLeuSerLys-----As 262
Qy 508 GGAGTTTATGAATGAATAAGCTGGGCTACGGCTC----- 541
Dy 262 nGluVal-----LeuArgLeuAlaMetLysTyrlleGlyPheLeuVa 276
Qy 542 ---CTGTTTGTAGCAGTGTGAGGAGTCTGGAGCA-----CCGATGGA----- 583
Dy 276 lArgLeuLeuArgAspGlnAlaAlaAlaLeuAlaGlyProThrProGlyProAr 296
Qy 584 -----CTCGAACGTGCTTTTCTCAAGTATGAAGACATGATCGGACCTGGTGA 633
Dy 296 gLysArgProValHisAsa:gvalProAspAspGlyProArgArgGlySerGlyArgArgAl 316
Qy 634 CGATGGTGGAGCAGCTGGCGCAGATTCTCTGGGGGTGTCTGTGTGACAAGGCCCGAGCTGGAAG 693
Dy 316 aGluAlaAlaAlaArgSerGln-----ProAlaProProAlaAs 329
Qy 694 CCTGACGGAGCACTGCCACCAGCTGTGTGGACCAAGTG 730
Dy 329 pProAspGlySerProGlyGlyAlaAlaArgProIle 341

RESULT 35
US-10-511-937-2466
; Sequence 2466, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511.937
; PRIOR FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2466
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2466

Alignment Scores:
Pred. No.: 0.855      Length: 362
Score: 93.00          Matches: 67
Percent Similarity: 30.4%      Conservative: 20
Best Local Similarity: 23.4%    Mismatches: 133
Query Match: 2.1%           Indels: 66
DB: 6                   Gaps: 12

US-10-768-158-1 (1-2419) x US-10-511-937-2466 (1-362)

Qy 765 TTCCCCCGCCCGCCAGCGGAGG-----CCTCAGCGTTGCGAGCACTGGTCCACCA----- 718
Dy 82 PheSerGlyProGluProGlyGlyAlaProGlnThrCysAlaLeuAlaProSerGluAla 101
Qy 717 -----GCTGGTGGCAGTGCT 703
Dy 102 SerGlyAlaGlnTyProProProGluThrLeuGlyAlaTyAlaGlyGly----- 119
Qy 702 CCGTGAGGCTTCCAGCTGGCGCTTGTACAGGACACCCCGAGAACTGCCAGCTGCT 643
Dy 120 ---ProGlyLeuValAlaGlyLeuLeuGlySerGluAspHisSerGlyTrpValArgPro 138
Qy 642 CCACCATCGTCACCGAGTCCGATGCTCTCATCTTGAGAAAAAGACGCTTCGAGT 583
Dy 139 AlaLeuArgAlaArgAlaProAspAlaPheVal-----Gly 150
Qy 582 CCATGCGGTGCTCCCAACTCTCTGCGAGCTCTCAACACGAGGCGGTAGCCAGCTTAT 523
Dy 151 ProAlaLeuAlaProAlaProAlaProGluProLysAlaLeuAlaGlnProValTy 170
Qy 522 -----CATTCTAAACCTCCGGCAGAACTTCTGAAAGGTGCTCGGTAGCTCATGG 472
Dy 171 ProGlyProGlyAlaGlySerSerGlyGlyTyPheProArg---ThrGlyLeuSerVal 189
Qy 471 TCCGACGAGCGGTGGAACTGATAATAAGACACCAACAGATCCTTGGGGTTGC----- 418
```

```

Db      190 ProAlaAlaSerGlyAla-----ProTyrGlyLeuLeuSer 201
Qy      417 ---GAGCCATATAGATGACCTTGGAGTCTCCATTGTGGAGGTTCAGAGGCGCAGAAAGCGGT 361
Db      202 GlyTyrProAlaMetTyrProAlaProGlnTyrGlnGlyHisPheGlnLeuPheArgGly 221
Qy      360 AGGGCAGGTGGCTCTTGATGAGGGGGGAGAGGTGATGCTTGTGATGATGTCAGGCCCG 301
Db      222 LeuGlnGlyProAla-----ProGlyPro 229
Qy      300 GCTGTGGGTACTCCA-----GGACCGGAGCTGCTCGTCGATGTTCACTCA 256
Db      230 AlaThrSerProSerPheLeuSerCysLeuGlyProGlyThrVal-----GlyThr 246
Qy      255 AGCCGATCTCATCGGGGTACAGCGCTCGCTCACCACCAAGTAGACCACTCTCTGCAGCAAGC 196
Db      247 GlyLeuGlyGlyThrAlaGluAspProGlyValIleAlaGluThrAlaProSerIysArg 266
Qy      195 TGGTCCCGGACTTGGGTAGGTAGCATCCACACGTCGCTGGCCGCA---CCGGGAAGT 139
Db      267 GlyArgArgSerTrpAlaArgIysArgGlnAlaAlaHisThrCysAlaHisProGlyCys 286
Qy      138 TGGCGATCTCTCCATCTTCCCGCGGAGAGGGCGGAGCGGCGGCGCATGCGAACTCGA 79
Db      287 GlyIysSerTyrThrLysSerSerHisLeuIysAlaHisLeuArgThrHisThrGlyGlu 306
Qy      78 AGTACTCTCTCGAAGTCCCGCGGGGTGCTGGGGGTCTCGGCTCTCGCTCTCCGCGCATGC 19
Db      307 LysProTyrAlaCysThrTrpGluGlyCysGlyTyrArgPheAlaArgSerAspGluLeu 326
Qy      18 CGCGCGCTCGCGCTCGC 1
Db      327 ThrArgHisTyrArgLys 332

RESULT 36
US-10-196-749-238
; Sequence 238, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430P1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; PRIOR FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541

```

```

; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION DATA removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 238
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-238

Alignment Scores:
Pred. No.: 0.903 Length: 325
Score: 92.50 Matches: 62
Percent Similarity: 32.0% Conservative: 24
Best Local Similarity: 23.0% Mismatches: 72
Query Match: 2.1% Indels: 111
DB: 6 Gaps: 16

US-10-768-158-1 (1-2419) x US-10-196-749-238 (1-325)
Qy      86 CCATGGCGTGGCGTGGCGGCTTCTGCGCGGGAAGATGGAGGATCGCAACTTCCC 145
Db      115 ProTrpGlnMetMetAlaAla-----
Qy      146 GGTGCGGCCAGCGACGTGTGGATGTCACCTACCCCAAGTCCGGCACCAGCTTGTGTCA 205
Db      122 ---AlaAlaGluAspArgLeuGlyGluAlaLeuProGluGluLeuSerTyrLeuSerSer 140
Qy      206 GGAGGTGGTCTACTTGTGTGGAGCCAGGGCGGTGACCCCGATGAGATCGCGTGTGATGAACAT 265
Db      141 AlaAlaAlaLeu-----
Qy      266 CGACGAGCAGTCCCGGTCTGTGGAGTACCCACAGCGCGGCTGGACATCATCAAGGAAC 325
Db      145 -----AlaProGlySerGly---ProLeuProGlyGluSerSerProAspAlaThr 160
Qy      326 GACCTCTCCCGCT-----CATCAAGAGCCACCTGCGCTACCGCTTCTTCTGCC 373
Db      161 GlyLeuSerProGluAlaSerLeuLeuHisGlnAspSerGluSerArgArgLeu----- 178
Qy      374 CTCTGACCTCCACAATGGAGACTCCAAGGTCTATATGCTGCGCAACCCCAAGGATCT 433
Db      179 -----ProArg-----SerAsnSerLeuGlyAla 186
Qy      434 GGTGGTGTCTATTATCAGTT---CCACCGCTCTCTCGGACCATGAG---CTACCCGAGG 487
Db      187 GlyGlyIysIleLeuSerGlnArgProTrpSerLeuIleHisArgValLeuProAsp 206
Qy      488 CACCTT-----
Db      207 HisProTrpGlyThrLeuAsnProSerValSerTrpGlyGlyGlyProGlyThrGly 226
Qy      494 -----TCAAGAATTCTCGCGAGGTTTATGAATGATAAGCTGGCTACGCTCTCTGTT 547
Db      227 TrpGlyThrArgProMetProHisPro-----GluGlyIle----- 238
Qy      548 TGAGCAGCTGCAGGAGTTCTGGGAGACCCGATCGGACTCGAAGCTGCTTTTCTTCAAGTA 607
Db      239 -----TrpGlyIleAsnAsnGlnProGly-----ThrSerTrpGlyAsnIle 253
Qy      608 TGAAGACATGATCGGAGCTGTGTGACGATGGTGTGAGCAGCTGCCAGATTCCTGGGGGT 667
Db      254 AsnArgTyr-----ProGlyGlySerTrpGlyAsnIleAsnArgTyrProGlyGly 270
Qy      668 GTC-----CTGTGACAAGGCCAGCTGGAGC-----
Db      271 SerTrpGlyAsnIleAsnArgTyrProGlyGlySerTrpGlyAsnIleHisLeuTyrPro 290
Qy      695 ---CCTGACGAGCACTGCCACCAAGCTGGTGGACAGTGTGCAACGCTGAGGCCCTGCC 751
Db      291 GlyIleAsnAsnProPheProPro-----GlyValLeu 301

```





```

; APPLICANT: Lee, Nancy M
; TITLE OF INVENTION: DRUG SCREENING AND MOLECULAR DIAGNOSTIC TEST FOR EARLY DETECTION
; FILE OF INVENTION: OF COLORECTAL CANCER: REAGENTS, METHODS, AND KITS THEREOF
; FILE REFERENCE: NLEE-01001US1 MCF/MLB
; CURRENT APPLICATION NUMBER: US/11/242,111
; CURRENT FILING DATE: 2005-09-29
; PRIOR APPLICATION NUMBER: 60/614,746
; PRIOR FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: 60/651,344
; PRIOR FILING DATE: 2005-02-08
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20
; LENGTH: 554
; TYPE: PRT
; ORGANISM: HUMAN
US-11-242-111-20

Alignment Scores:
Pred. No.: 1.35 Length: 554
Score: 91.50 Matches: 74
Percent Similarity: 30.5% Conservative: 22
Best Local Similarity: 23.5% Mismatches: 129
Query Match: 2.1% Indels: 90
DB: 7 Gaps: 18

US-10-768-158-1 (1-2419) x US-11-242-111-20 (1-554)
QY 832 TGTATTATACCAAGTCAAACTTCTCATTCATGACGAGCGGTGAAGATGCTCTCCACAGC 773
DB 191 CysleuYrProLysAla---IleProSerSerAspProAlaSerValSerProHisGln 209
QY 772 CCAACTCTTCCCGGCCCCACGCGGAGCCCTCAGCGTTGCGACACTGCTGCCACAGCTGG 713
DB 210 ProLeuAlaProSerMetAlaProValAlaGlyLeuThr-----TIpGluAspSer-Gl 227
QY 712 TGGCAGTCTCGGTCAGGCTCCAGCTGG-----CCTGTGCACAGCACACCCGAGG 659
DB 227 uGlyThrGluGlySerSerLeuLeuProGlyGluGlnProLeuHisThrValAspProGl 247
QY 658 AATCTGGCCAGCTGCTCCACCATCTCA-----CCAGGTCGCCGATGCATG 614
DB 247 ySerAlaLysGlnArgProProArgSerThrCysGlnSerPheGluProProGluThrPr 267
QY 613 TCTTATATCTTGAGAAAAAGCAGCGTTCGAGTCCATGCGGTGCTCCAGAACTCTGACAG 554
DB 267 oVal-----ValLysAspSerThrIleGlyGlySerProGlnPro--Ar 281
QY 553 TGCTCAACACGAGGCGGTAGCCAGCTTATCATTTCAATAACCTCCGCGAGATTCTTGA 494
DB 281 gProSerValGlyAlaPheAsnProGlyMetGluAsp-----IleLeuAs 296
QY 493 AAGGTGCTCGGTAGCTCATGTCGCCGAGAGCGGTGGAACCTGATAATAAGACACACC 434
DB 296 pSerAlaMetGlyThrAsnTrpValProGluGluLaserGlyGluAlaSerGluIlePr 316
QY 433 AGATCCTTGGGGTTCGAG-----CCATATAGATGACCTTGGAGTCTCCATTTGGGAGG 380
DB 316 oValProGlnGlyThrGluLeuSerProSerArg--ProGlyGly-----Glyse 332
QY 379 TCAGAGGCGCAGAACCGGTAGGCGAGGTGGCTCTTGATGAGCGGGGAGAGGTTCAGTCC 320
DB 332 rMetGlnThrGluProAlaArgProSerAsnPhe-----LeuSerAlaSerSerPr 349
QY 319 TTGATGATGTCAGGCGCGGTGTGGGT-----AC 290
DB 349 o-----LeuProAlaSerAlaLysGlyGlnGlnProAlaAspValThrAlaThrAlaLe 367
QY 289 TCCAGGACCGGAGGTGCTCGTCGATGTTTCATCAAGCCGATCTCATCGGGGTACGCGCC 230
DB 367 uProArgValGlyProValMetProThrGlyGlnAspTrpAsnHisThrProGlnLysTh 387
QY 229 TGGCTCACCAAGT-----AGACCACCTCTCTGCAGCAAGCTGCTGCCGAGTGGGG 179
```

```

DB 387 rAspHisProSerAlaLeuLeuArgAspProPro----- 398
QY 178 TAGGTGACGATCCACACGTCGCTGGCGCCGACCGGGAAGTTGGCGA----- 133
DB 399 -----GluProGlySerProArgIleSerSerLeuAr 409
QY 132 -----TCTCTCCATCTTCCCGGGCAGAGGGCGGCACCGCACGC----- 91
DB 409 gProGlnAlaLeuSerAsnProSerThrLeuSerAlaGlnProGlnLeuSerArgSerHi 429
QY 90 -----CATGGAACCTCAAAGTACTTCTCTCTCG 65
DB 429 sSerSerGlySerValLeuProLeuGlyGluLeuGluGlyArgArgSerThrArgAspAr 449
QY 64 AACTCCCCCGGGGTGCTGGGGGTCTCGGCTCTCGCTCTCCGCCA 22
DB 449 gThrSerPro-----AlaGluProGluAlaAlaPro 459

RESULT 41
US-10-322-836-48
; Sequence 48, Application US/10322836
; Publication No. US2006090212A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Alexander C., Jr.
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Freidrich, Glenn A.
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND PROTEINS
; TITLE OF INVENTION: ENCODED THEREBY
; FILE REFERENCE: 8535-0037-999
; CURRENT APPLICATION NUMBER: US/10/322,836
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: US/09/579,114
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 1075
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (130)...(130)
; OTHER INFORMATION: Xaa = Gln or STOP
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (179)...(179)
; OTHER INFORMATION: Xaa = Ser or Gly
US-10-322-836-48
```

```

Alignment Scores:
Pred. No.: 1.74 Length: 1075
Score: 91.50 Matches: 74
Percent Similarity: 31.9% Conservative: 23
Best Local Similarity: 24.3% Mismatches: 89
Query Match: 2.1% Indels: 118
DB: 6 Gaps: 16
```

```

US-10-768-158-1 (1-2419) x US-10-322-836-48 (1-1075)
QY 768 CTCCTCCCGGCCACCG-----GCAGGGCT-----CAGCGTTCGACACTGGTCCA 721
DB 669 IlePheProSerProArgGluLeuGluGlyProValTyGluLysCysMet----- 685
QY 720 CCAGCTGTGGCAGTGTCTCGCTCAGGGCTTCCAGCTGGCGCTTGTCTCAGGACACCCCA 661
DB 686 ---AlaGlyGlyGluGluTyCysAspSerPro-----HisSerGluProGly 700
QY 660 GGAATCTGGCAGCTGTCTCCACCATCGCTCACAGGTCCCGATGCTTCTTCACTATGTA 601
DB 701 AlaIle-AspGluValaAspHisAsnGlyThrGluPro---HisThrSerAspGluGl 719
```







```
DB: 6 Gaps: 10
US-10-768-158-1 (1-2419) x US-10-509-131-25 (1-783)
QY 786 TGCTCTCCACAGCCAACTCTCCCGCCACCGGCGAGG-----CCT 742
Db 633 CysProSerThrSerSerLeuPheAsnLeuGlnLysSerSerLeuSerAlaArgHisPro 652
QY 741 CAGCCTTGACGACTGGT-----CCACGAGCTGGTGGCAGTGTCCGTCAGGGCTT 691
Db 653 GlnArgLysArgArgGlyGlyProSerGluProThrProGlySerArgProGlnAspAla 672
QY 690 CCAGCTGGCCTTGT-----CCACGACACCCCGCAGGATCTGGCCAGCT 646
Db 673 ThrValHisProAlaCysGlnPheProHisTyrThrProSerValAlaTyrProTrp 692
QY 645 GCTCCA-----CCT 637
Db 693 SerProGluAlaHisProLeuLeuCysGlyProProGlyLeuAspLysArgLeuLeuPro 712
QY 636 TGTCTCAGGCTCCGATGTCATGTCTTTCATCTTGAGAAAGACGTTCCGAGTCCA--- 580
Db 713 GluThrProGlyPro-----CysTyr-----SerAsnSerGlnProVal 725
QY 579 -----TGGGTGCTCCAGAACTCTCGACGTCGTCTAAACGAGGCGGTAGCCAGCT 526
Db 726 TrpLeuCysLeuThrProArgGlnProLeuGluProHisProProGlyGluGlyPro--- 744
QY 525 TATCATTCAAACTCCGGCAGATTCTTGAAAGGTGCTTGGGGTTCGCGAGCCATAGA 466
Db 745 -----AlaGluGlyArgProCysProTyrPro-----SerGluTrpSerSer 749
QY 465 GAGAGCGGTGGAAGTATAAAGACACACAGATCCTTGGGGTTCGCGAGCCATAGA 406
Db 750 AspThr-----AlaGluGlyArgProCysProTyrPro-----SerGluTrpSerSer 749
QY 405 TGACCTGGAGTCTCCATTGGGAGGTGAGAGGTGAGAGGCGAGAAAGCGGT 361
Db 761 -----HisCysGlnValLeuSerAlaGlnProGly 770
RESULT 45
US-10-196-749-236
; Sequence 236, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; PRIOR FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
```

```
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 236
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-236
Alignment Scores: 1.38 Length: 284
Pred. No.: 50.00 Matches: 48
Score: 27.6% Conservative: 11
Percent Similarity: 22.4% Mismatches: 53
Best Local Similarity: 2.0% Indels: 102
Query Match: 6 Gaps: 9
DB:
US-10-768-158-1 (1-2419) x US-10-196-749-236 (1-284)
QY 810 TCTCATTATGAGAGCGGTGAAGATGCTCTTCCACAGCCCAACTCTTCCCGGCGCCACGG 751
Db 3 SerTyrProTyrArgGln---GlyCysProGlyAlaGlyGlnAlaProGly----- 19
QY 750 GCAGGGCTCAGCGTTCGAGCACTGGTCCACCA---GCTGGTGGCAGTGTCCGTCAGGG 694
Db 20 AlaProProGlySerTyrTyrProGlyProProAsnSerGlyGlyGlnTyrGlySerGly 39
QY 593 CTTCAGCTGGCGCTTGTTCACAGACACCCCGCAGGATCTCCGCGAGTGTCCACCATCG 634
Db 40 LeuProProGlyGlyGlyTyrGlyGlyProAlaProGlyGlyProTyrGlyProProAla 59
QY 633 TCACGAGTCCCGATGCATGTCTTCTATCTTGAGAAAAAGCACGTTCCAGTCCATCGGT 574
Db 60 GlyGlyGly-----ProTyrGly 65
QY 573 GCTCCAGAACTCCTGCACGTGCTCAAACAGGAGCGGTAGCCAGCTTATCATTTATAA 514
Db 66 HisPro----- 67
QY 513 ACCTCCGCGCAATTCTTGAAGGTGCTCGGTAGCTCATGTCGCGAGAGCGGTGGA 454
Db 68 -----AsnProGlyMet 71
QY 453 ACTGATAATAAGACACACACAGATCCTTGGGGTTCGCGAGCCATATAGATGACCTTGGAGT 394
Db 72 PheProSerGlyThrProGlyGlyProTyrGly----- 82
QY 393 CTCCATTGTGGAGTTCAGAGGCGAGAAAGCGGTAGGCGAGGTGGCTCTTGTAGAGCGGG 334
Db 83 -----GlyAlaAla----- 85
QY 333 GAGAGGTCACTTCTTGTATGTTCAGGCGCGGCTGTGGGTACTCCAGGACCGGGAGCT 274
Db 86 -----ProGlyGlyProTyrGlyGlnProProSer--- 96
QY 273 GCTCGTCGATGTTTCATCAAGCCGATCATCGGGGTGAGCGCGCTCGG----- 226
Db 97 -----SerTyrGlyAlaGlnProGlyLeuTyrGlyGln 108
QY 225 -----TCACCAAGTAGACACCATCTCTTCGACCAAGCTGG 193
Db 109 GlyGlyAlaProProAsnValAspProGluAlaTyrSerTrp 122
RESULT 46
US-11-101-316-62
; Sequence 62, Application US/11101316
```

```

; Publication No. US20060099657A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
; TITLE OF INVENTION: UNDEREXPRESSED IN MELANOMA
; FILE REFERENCE: P3230R1C17C1
; CURRENT APPLICATION NUMBER: US/11/101,316
; CURRENT FILING DATE: 2005-04-06
; PRIOR APPLICATION NUMBER: 10/063526
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 62
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-11-101-316-62

Alignment Scores:
Pred. No.: 1.38 Length: 284
Score: 90.00 Matches: 48
Percent Similarity: 27.6% Conservative: 11
Best Local Similarity: 22.4% Mismatches: 53
Query Match: 2.0% Indels: 102
DB: 9 Gaps: 9

US-10-768-158-1 (1-2419) x US-11-101-316-62 (1-284)

Qy 810 TCTCATTCATGAGACGGTGAAGATCTCTTCACAGCCCACTTCTCCCGGCCACGG 751
Db 3 SerTyProTyArgGln--GlyCysProGlyAlaAlaGlyGlnAlaProGly----- 19
Qy 750 GCAGGGCTCAGGTTGCAGCACTGCTCCACCA--GCTGGTGGCAGTGCTCCGTCAGG 694
Db 20 AlaProProGlySerTyTyProGlyProGlyProAlaProGlyGlyGlnTyGlySerGly 39
Qy 693 CTTCCAGCTGGGCTTGTACAGACGACCCCGAGGAATCTGCCCACTCTCCACCATCG 634
Db 40 LeuProProGlyGlyTyGlyTyGlyTyGlyTyGlyTyGlyTyGlyTyGlyTyGlyTyGly 59
Qy 633 TCACCAAGTCCCGATGATGCTTCTTCATCTTGAGAAAAAGCAGTTCGAGTCGCGT 574
Db 60 GlyGlyGly-----ProTyGly 65
Qy 573 GCTCCCAAGAACTCTGCACGTGCTCAACACGAGGCGGTAGCCAGCTTATCATTTAA 514
Db 66 HisPro----- 67
Qy 513 ACCTCCGGCAGAAATCTTGAAGAGTGCCTCGGTAGCTCATGTCGAGAGCGGTGA 454
Db 68 -----AsnProGlyMet 71
Qy 453 ACTGATAATAAGACACCAACAGATCCTTGGGGTTCGAGCCATATAGATGACCTTGGAGT 394
Db 72 PheProSerGlyThrProGlyGlyProTyGly----- 82
Qy 393 CTCATTGTGGAGTTCAGAGGCGAGAAACGGGTAGGCGAGGTGGCTCTTGTAGAGCGGG 334
Db 83 -----GlyAlaAla----- 85

```

```
Db 28 LeuSerCysThrTyrLeuCysTyrSerPheLeuCysCysCys-----AspAspLeuGly 45
Qy 121 AGATGGAGAGATCGCCAACTTCCCGGTGC-----GCCCGAGCGACGTGTGGA----- 168
Db 46 ArgSerArgLeuLeuGlyAlaProArgCysLeuArgGlyProSerAlaGlyGlyGlnLys 65
Qy 169 -----TCGTCACCTACCCCAAGT----- 186
Db 66 LeuLeuGlnLysSerArgProCysAspProSerGlyProThrProSerGluProSerAla 85
Qy 187 -----CCGGACACAGCTTGTGCAGGAGGTGCTACTTGGTAGCCAGGCGGTG 237
Db 86 ProSerAlaProAlaAlaAla-ValProAlaProArgLeuSerGly-SerAsnHisSerG 105
Qy 238 ACCCGATGAGATCGGCTTGATGAACATCGACGACGCTC----- 278
Db 105 lySerProLysLeuGlyThrLysArgLeuProGlnAlaLeuIleValGlyValLysLysG 125
Qy 279 -----CCGGTCTCGAGTAC-----C 294
Db 125 lyGlyThrArgAlaValLeuGluPheIleArgValHisProAspValArgAlaLeuGlyT 145
Qy 295 CACAGCG-----GCCCTGGACATCATCAAGNACTGA 327
Db 145 hrGluProHisPhePheAspArgAenTyrGlyArgGlyLeuAspTrpTyrArgSerLeu- 164
Qy 328 CTTCTCCCGCTCATCAAGACCACTCGCCCTACCGCTTCTGCGCTCTGCACCTC----- 383
Db 165 --MetProArgThrLeuGluSerGlnIleThrLeuGluLysThrProSerTyrPheValT 184
Qy 384 -----CACAAATGGAGACTCCAAGGTCTATATATGG 414
Db 184 hrGlnGluAlaProArgArgIlePheAsnMetSerArgAspThrLysLeuIleValValV 204
Qy 415 CTGCAACCCCAAGATCTGTGTGTCTTATTATTCAGTTCACCGCTCTCTCGGACCA 474
Db 204 alArgAsnProValThrArgAlaIleSerAspTyr-----ThrGlnThrL 219
Qy 475 TGAGCTACCGA-----GGCACCTTTCAAGAAATCTCGCGGAGGTTTATGAATGATA 525
Db 219 euSerLysLysProAspIleProThrPheGluGlyLeu-----SerPheArgAsnArgT 237
Qy 526 AGCTGGGC-----TACGGCTCTCGTGTGGACGACG 555
Db 237 hrLeuGlyLeuValAspValSerTrpAsnAlaIleArgIleGlyMetTyrValLeuHisL 257
Qy 556 TGCAGGAGTTCTGGGAGCACCGCATCGACTCGAAGCTGCTTTTCTCAAGTATGAAGACA 615
Db 257 euGluSerTrpLeuGlnTyrPheProLeuAlaGlnIlePheValSerGlyGluArgL 277
Qy 616 TGCATCGGACCTGCTGACGATGTGTGGAGCAGCTGGCCAGATTCTCGGGGTG 668
Db 277 euIleThrAspProAlaGlyGluMetGlyArgValGlnAspPheLeuGlyIle 294
```

## RESULT 48

```
US-10-505-928-438
; Sequence 438, Application US/10505928
; Publication No. US20060089532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 438
; LENGTH: 717
; TYPE: PRT
; ORGANISM: Homo sapiens
```

## US-10-505-928-438

```
Alignment Scores:
Pred. No.: 1.99 Length: 717
Score: 90.00 Matches: 69
Percent Similarity: 32.9% Conservative: 24
Best Local Similarity: 24.4% Mismatches: 89
Query Match: 2.0% Indels: 101
DB: 6 Gaps: 17
```

US-10-768-158-1 (1-2419) x US-10-505-928-438 (1-717)

```
Qy 37 CCGAGACCCACAGCACCCCGGGGAGTTCGAGACAAGTACTTCGAGTTCCATGCGCTGC 96
Db 299 ProGlyProProGlyProLysGlyAspProGlyIleGln----- 311
Qy 97 GGCTGCCGCCCTTCTGCCGGGAGATCGGAGAGATCGCAACTTCCCGGT----- 148
Db 312 -----GlyTyrHisGlyArg-LysGlyGluArgGlyMetProGlyMetProG 327
Qy 149 -----GCCGCCACGACGCTGTGGATCGTCA 174
Db 327 yLysHisGlyAlaLysGlyAlaProGlyIleAlaValAlaGlyMetLysGlyGluProG 347
Qy 175 CTTACCCCAAGTCGGCACACAGCTTGTGTGAGGAGGT-----GGTCTACT 219
Db 347 yIleProGlyThrLysGlyGluLysGlyAlaGluGlySerProGlyLeuProGlyLeu 367
Qy 220 TGGTGAGCCAGGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGA----- 268
Db 367 uGlyGlnLysGly-----GluLysGlyAspAlaGlyAsnSerIleGlyG 382
Qy 269 ----CGAGCAGCTCCCGGT---CCTGGAGTACCC---ACAGCCGGCGCTCGACATCA 318
Db 382 yGlyArgGlyGluProGlyProProGlyLeuProGlyProProGlyProLysGlyGlu 402
Qy 319 AGGAACCTGAC-----CTCTCCCGCTCATCAAGAGCCACCTGCCCTTCTGCG 372
Db 402 aGlyValAspGlyGlnValGlyProProGlyGln----- 413
Qy 373 CCTCTGACCTCCCAATGGAGACTCCCAAGGTCTATATGGCTCGCAA----- 421
Db 414 -----ProGlyAspLysGlyGluArgGlyAlaAlaGlyGlnGlyProAspG 430
Qy 422 -CCCCAAGATCTGCTGTGTCTTATTATCAGTTCCACCGCTCTCTGCGGACCATGAGCT 480
Db 430 yProLysGlySerLysGly-----GluProGlyLysGlyGluMetValAspTy 446
Qy 481 ACCGAGCAGCTTTCAGAAGATTCGCGGAGGTTTATGAATGATAAGCTGGGTACGGCT 540
Db 446 rAsnGlyAsnIleAsnGluAlaLeuGlnGluIle----- 457
Qy 541 CCTGTTTGACACGCTGAGGAGTTCCTGGGAGCACCGCATGAGCTCGAACGCTGCTTTTC 600
Db 458 -----ArgThrLeuAlaLeuMetGlyProProGlyLeuPro-----G 470
Qy 601 TCAAGTATGAAGACATGCATCGGACCTGTGTGACGATGGT-----GGAGCAGCTGGCCA 654
Db 470 yGlnIle-----GlyProGlyAlaProGlyIleProGlyGlnLysGlyG 486
Qy 655 GATT-----CCTGGGGGTGCTCTGTGACAGGCCAGCTGGAAGCCCTGA 699
Db 486 uIleGlyLeuProGlyProProGlyHisAspGlyGluLysGlyProArgGlyLysProG 506
Qy 700 CGGAGCACTGCCACCACTGCTGGACCACTGCTGCACGCTGAGGCGCTGCCCGTGGGCC 759
Db 506 yAspMetGlyProPro-----GlyProGlnGlyProPr 517
Qy 760 GGGGAAG 766
Db 517 oGlyLys 519
```

RESULT 49

```
US-10-505-928-493
; Sequence 493, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 493
; LENGTH: 744
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-493

Alignment Scores:
Pred. No.: 2.22 Length: 744
Score: 89.50 Matches: 75
Percent Similarity: 25.4% Conservative: 11
Best Local Similarity: 22.1% Mismatches: 76
Query Match: 2.0% Indels: 177
DB: 6 Gaps: 17

US-10-768-158-1 (1-2419) x US-10-505-928-493 (1-744)
QY 783 CTTTCCACAGCCCAACTCTCCCGGCCACGG-----GCAGGGCCTCAG 739
DB 403 ProProGlyAlaIleGlyPheProGlyProLysGlyGluGlyIleValGlyProGln 422
QY 738 CGTTGCACACTGTCACACAGCT----- 715
DB 423 -----GlyProProGlyProLysGlyGluProGlyLeuGlnGlyPheProGly 438
QY 714 -----GGTGGCAGTGTCCGTCACGGCTT-----CCAGCT 685
DB 439 LysProGlyPheLeuGlyGluValGlyProGlyMetArgGlyLeuProGlyProIle 458
QY 684 GGGCCTTGTACAGGACA----- 667
DB 459 GlyProLysGlyGluAlaGlyGlnLysGlyValProGlyLeuProGlyValProGlyLeu 478
QY 666 -----CCCCCAGGA 658
DB 479 LeuGlyProLysGlyGluProGlyIleProGlyAaspGlnGlyLeuGlnGlyProProGly 498
QY 657 ATC-----TGGCCAGCTGTCCACCATCGTCACACAGGTCCCGAT 619
DB 499 IleProGlyIleGlyGlyProSerGlyProIleGlyProGlyIleProGlyPro---- 517
QY 618 GCATGCTTCATCTTGAGAAAGACAGCTTCGAGTCCATCGGTGCTCCCGAACACTCCT 559
DB 518 -----LysGlyGluProGlyLeuProGlyProGlyPhePro 530
QY 558 GCACGTGCTCAACACAGGAGCGTAGCCAGCTTATCATTAATCAATAACCTCCGGCAGATT 499
DB 531 Gly-IleGlyLysProGlyValAlaGly-----LeuHisGlyProPro----- 544
QY 498 CTTGAAGGTGCCTCGGTAGCTCATGTGTCGACAGAGCGGTGGAACTGATAATAAGACA 439
DB 545 -----GlyLysProGlyAlaLeuGlyProGln----- 553
QY 438 CCACAGATCCTTGGGGTTGGAGGCATATAGATGACCTTGGAGTCTCCATTGTGGAGGT 379
DB 553 ----- 553
QY 378 CAGAGGGCAGAAAGCGGTAGGCGAGGTGGCTTTGATGAGCGGGGAGAGGTTCCTT 319
DB 554 ----GlyGln--ProGlyLeuProGlyPro----- 561
```

```
318 TGATGATGTCAGGCCGCGGTGTGGGTACTCCAGACCGGAGCTGCTCGTGATTTCA 259
DB 562 -----ProGlyProPro-----GlyProProGlyProProAlaVal----- 573
QY 258 TCAAGCCGATCTCATCGGGGTACAGCCCTCGGTTCACCAAGTAGACACCTCTCTGAG-- 201
DB 574 -----MetProProThrProProPro-GlnG1 582
QY 200 -----CAAGCTGTGCCGACTTGGGTAGGTGACGATCCACACCTCGCTGG 154
DB 582 yGluTyLeuProAaspMetGlyLeuGlyIleAaspGlyVallyProProHisAlaTyG1 602
QY 153 GCGGCAC-----CGGGAAGTTGGCGATCTCTCCATCTTCCCGCG-----GCAGAA-- 108
DB 602 yAlaLysGlyLysAaspGlyProAlaTyGluMetProAlaPheThrAlaGluLe 622
QY 107 ----- 106
DB 622 uThrAlaProPheProProValGlyAlaProVallyPheAaspLysLeuLeuTyAaspG1 642
QY 105 GCGGCAGCCGACGCCATGGAACCTGGAAGTACTTGTCTCTCGAACTCCCC 57
DB 642 yArgGlnAaspTyAsnProGlnThrGlyIlePheThrCysGluValPro 658

RESULT 50
US-10-505-928-537
; Sequence 537, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 537
; LENGTH: 1193
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-537

Alignment Scores:
Pred. No.: 2.94 Length: 1193
Score: 89.00 Matches: 97
Percent Similarity: 34.3% Conservative: 41
Best Local Similarity: 24.1% Mismatches: 136
Query Match: 2.0% Indels: 128
DB: 6 Gaps: 24

US-10-768-158-1 (1-2419) x US-10-505-928-537 (1-1193)
QY 1125 CCAGGCCAT-----TGGAACTGCAATGT-----GAGACTGTTTGTAAAT 1087
DB 463 ProCysHisAaspGlyPheSerCysSerValMetProGluThrGluGluValValCysAan 482
QY 1086 CAGACATGGAGAGCTGCAGCTTCTAAAGGCGAGACAGCTGC----- 1045
DB 483 AsnCysProProGlyValThrGlyAlaArgCysGluLeuCysAlaAaspGlyTyxPheGly 502
QY 1044 -----TTCGGTTGGGAATCATCACTCCCTCCGCTCACGCCCTCTTCCCTTCCCGCG 991
DB 503 AspProPheGly-----GluHisGly--ProValArgProCysGlnPro----- 516
QY 990 CTGTTTCACAGCTGCTTCCAGAGTTTGTCCAGAAAGATAAATAATGATCATACAGGAC 931
DB 517 -----CysGlnCysAaspAaspAan----- 522
QY 930 TTTTGGCTAGTAGACTGTCTGGGTATTGTGAGCAT-----GCAGGTTGTTGTTCTGTT 877
DB 523 --ValAaspProSerAlaSerGlyAaspCysAaspArgLeuThrGlyArgCysLeuLysCys- 541
```



```
QY 132 TCTCTCCATCTTCCCGGCGAGAGGGCGGCGCAGCCGCGCAGCCATCGAAGTACT 73
Db 227 -----ProArgAspAlaGly-----AsnGlyProGlnA 236
QY 72 TGCTCTCGAACTCCCGCGGGGTG---CTGGGGGTCTCGGCTCGCTCTCCGCCATGCCGC 16
Db 236 laLeuArgGlnAlaSerGlyAlaProGlnGlyGluAlaAlaProGluThrAspSerSerp 256
QY 15 CGCGCTCGCGCTCG 2
Db 256 roProSerProSer 260

RESULT 52
US-10-505-928-831
; Sequence 831, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-831

Alignment Scores:
Pred. No.: 3.98 Length: 2026
Score: 88.50 Matches: 68
Percent Similarity: 27.3% Conservative: 13
Best Local Similarity: 22.9% Mismatches: 121
Query Match: 2.0% Indels: 95
DB: 6 Gaps: 13

US-10-768-158-1 (1-2419) x US-10-505-928-831 (1-2026)
QY 786 TGCTCTCCAGCCGCACTCTCCCG-----GGCCCGAGG 751
Db 770 CysProSerAlaCysGlnCysAsnProHiesGlySerTyrGlyThrCysAspProAla 789
QY 750 GCAGGGCTCAGCGTTCGAGCACTGTCCACCACTGTGTGGCAGTGTCTCGTCAGGGCTT 691
Db 790 ThrGly-----GlnCysSerCysArgProGlyValGlyLeuArg----- 803
QY 690 CCAGCTGGGCTTCTCAGCAGACACCCCGAGAACTCTGGCCAGCTGTCTCCACCATCGTCA 631
Db 804 -----CysAspArgCysGluProGlyPheTrpAsnPheArgGlyIleValThr 819
QY 630 CCAGTCCGATGATGCTTCTCATCTACTTGAGAAAAGCAGCTTCGAGTCCATGC----- 577
Db 820 AspGlyArgSerGlyCys-----ThrProCysSerCys 830
QY 576 -----GGTGCTCCAGAACTCTCTCAGCTGTCAAACAGG-----AGCCGTAGC 532
Db 831 AspProGlnGlyAlaValArgAspAspCysGluGlnMetThrGlyLeuCysSerCysLys 850
QY 531 CCAGCTTATCATTAACCTCCCGCAGAAATCTTGAA-----AGTGCTCGGTAGCTCA 475
Db 851 ProGlyValAlaGlyProLysCysGlyGlnCysProAspGlyArgAlaLeuGlyProAla 870
QY 474 TGGTCCGAGAGCGGTGGAACTGATAATAAGACACCAACCCAGATCTCTGGGGTTGCAG 415
Db 871 GlyCysGluAlaAspAlaSerAla-----ProAlaThrCysAla 883
QY 414 CCATATAGATGACCTTGGAGTCTCCATTGTGGAGTTCAGAGGGCAGAAAGGGTAGGGCA 355
Db 414 -----ProAlaThrCysAla 883
QY 414 CCATATAGATGACCTTGGAGTCTCCATTGTGGAGTTCAGAGGGCAGAAAGGGTAGGGCA 355
Db 414 -----ProAlaThrCysAla 883
```

```
Db 884 GluMetArgCysGluPheGlyAlaArgCys-----ValGluGluSerGlySerAla 900
QY 354 GGTGGCTCTTCATGAGCGGGGAGAGGTCAAGTCTTCATGATGATGTCAGGCGGCTGTG 295
Db 901 His-----CysValCysProMetLeu 907
QY 294 GGTACTCCAGGACCGGAGCTGCTCGTGCATGTTTCATCAAGCCGATCTCATCGGGGT--- 238
Db 908 ThrCysProGluAlaAsnAlaThrLysValCysGlySerAspGlyValThrTyrGlyAsn 927
QY 238 ----- 238
Db 928 GluCysGlnLeuLysThrIleAlaCysArgGlnGlyLeuGlnIleSerIleGlnSerLeu 947
QY 237 -----CAGCGCCCTGGCTCACAAAGTAGACACCTCTCTCGCAGCAAGC---TGTGTC 190
Db 948 GlyProCysGlnGluAlaValAlaProSerThrHisProThrSerAlaSerValThrVal 967
QY 189 CGGACTTGGGGTAGGTGACGATCCACACGTCGCTGGGCGCGCACGGGAAGT----- 139
Db 968 ThrThrProGlyLeuLeuSerGlnAlaLeuProAlaProGlyAlaLeuProLeu 987
QY 138 -----TGGCGATCTCTCTCATCTCTCCATCTCTCCGCGGCGAGAGG 106
Db 988 AlaProSerSerThrAlaHisSerGlnThrThrProProProSerSerArgProArgThr 1007
QY 105 GCGGCGAGCGCAGCCCATGGAAGTCTGCTCTCGAACTCCCGG 55
Db 1008 ThrAlaSerValProArgThrThrValTrpProValLeuThrValProPro 1024

RESULT 53
US-10-505-928-805
; Sequence 805, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 805
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-805

Alignment Scores:
Pred. No.: 1.99 Length: 270
Score: 88.00 Matches: 51
Percent Similarity: 31.1% Conservative: 14
Best Local Similarity: 24.4% Mismatches: 87
Query Match: 2.0% Indels: 58
DB: 6 Gaps: 8

US-10-768-158-1 (1-2419) x US-10-505-928-805 (1-270)
QY 699 TCAGGGCTTCAGCTGGGCTTGTTCACAGACACCCCGAGAACTCGCCAGCTGCTCCA 640
Db 18 AsnGlyLeuHisHisHisProAlaHisArgMetGlyMetGlyGlnPheProSerPro-Hi 37
QY 639 C---CATCGTCACAGGTCGCGATGTCATGCTTC---ATACTTGAGAAAAGCAGCTTCG 586
Db 37 sHisHisGlnGlnGlnProGlnHisAlaPheAsnAlaLeuMetGlyGluHisIleHi 57
QY 585 AGTCCATGCG-----GTGCTCCAGAACTCTCTCGACGCTGTCTCAAACAGGAGC 538
Db 57 sTyrGlyAlaGlyAsnMetAsnAlaThrSerGlyIleArgHisAlaMetGlyProGlyTh 77
QY 537 CGTAGCCCGCTTATCATTAACCTCCCGGAGAAATCTTGAAGAGTGCC----- 486
```

Db 77 rValAsnGly-----GlyHisProProSerAlaLeuAlaProAlaAlaArgPheAs 94  
Qy 485 -----TCGGTAGCTCATGCTCG-----GlyHisProProSerAlaLeuAlaProAlaAlaArgPheAs 94  
Db 94 nAsnSerGlnPheMetGlyProProValAlaSerGlnGlySerLeuProAlaSerMe 114  
Qy 467 -----CAGAGAGCGGTGGAACTGATTAAGACAC----- 438  
Db 114 tGlnLeuGlnLysLeuAsnAsnGlnTyrPheAsnHisHisProTyrProHisAsnHisTy 134  
Qy 437 -----CACCAGATCCTTGGGTGGGTTGGAGCCATATAGA 406  
Db 134 rMetProAspLeuHisProAlaAlaGlyHisGlnMetAsnGlyThrAsnGlnHisPheAr 154  
Qy 405 TGACCTTGAGCT-CTCCATTGTGGAGGTTCAGAGGCGCAGAAAGCGGTAGGCGGTGGCTC 347  
Db 154 gAspCysAsnProLysHisSerGlyGlySerThrProGlyGlySerGlyGlySerSe 174  
Qy 346 TTGATGAGCGGGGAGAGGTGAGTTCCTTGATGATGTCAGGCGCGGTGGGTGACTCC 287  
Db 174 r-----ThrProGlyGlySerGlySerSe 183  
Qy 286 AGGACCGGAGCTGCTCGTCTGATGTTTCATCAAGCGATCTCATCGGGGTACGCGCTGG 227  
Db 183 rGlyGlyGlyAlaGlySerAsnSerGlyGlyGlySerGlyGlySerGlyGlySerProAl 203  
Qy 226 CTCACCAAGTAGACCACCTCTCTGCA 202  
Db 203 aSerValAlaHisValProAlaAla 211

RESULT 54

US-10-861-934-16  
; Sequence 16, Application US/10861934  
; Publication No. US20060089491A2  
; GENERAL INFORMATION:  
; APPLICANT: Nagata, Shigikazu  
; Suda, Takashi  
; Takahashi, Tomoniro  
; Nakamura, Norio  
; TITLE OF INVENTION: A Fas Ligand, A Fragment Thereof and DNA  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
; STREET: P.O. Box 747  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22040-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/861,934  
; FILING DATE: 07-Jun-2004  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/683,703  
; FILING DATE: 07-JUL-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Murphy Jr., Gerald M.  
; REGISTRATION NUMBER: 28,977  
; REFERENCE/DOCKET NUMBER: 1110-163P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-205-8000  
; TELEFAX: 703-205-8050  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 278 amino acids  
; TYPE: amino acid

; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-10-861-934-16  
Alignment Scores:  
Pred. No.: 2.01 Length: 278  
Score: 88.00 Matches: 61  
Percent Similarity: 34.5% Conservative: 26  
Best Local Similarity: 24.2% Mismatch: 72  
Query Match: 2.0% Indels: 93  
DB: 6 Gaps: 12  
US-10-768-158-1 (1-2419) x US-10-861-934-16 (1-278)  
Qy 963 GTCCAGCAAGGAATAAAT---GAATGCATACAGGACTTTTGGCTAGTAGACTGTCTGGGT 907  
Db 1 MetGlnGlnProValAsnTyrProCysProGlnIleTyrTrp----- 14  
Qy 906 ATTGTGAGCATGCAGGTGTGTTTCTGTATTATAATAAAAGTCAAAAGTCAAGGTGCAC 847  
Db 15 -----ValAspSerSerAlaThrSerProTrp 23  
Qy 846 ACTTTCCTCATCTTCTGTTTATACACCAAGTCAAACTTCTCATTATGAGAGCGGTGAAGA 787  
Db 24 AlaProProGlySerValPheSer----- 31  
Qy 786 TGCTCTTCCACAGCCCAACTCTTCCCGGCGCCACGGCAGGCGCTCAGCGTTGCAGCACT 727  
Db 32 CysProSer-----GlyProArgGlyProGlyGlnArg----- 43  
Qy 726 GGTCACCAAGCTGTGGCAGTGTCTCCAGGGGTTCACAGCTGGGCTTGTTCACAGGACA 667  
Db 44 ArgProProProProProProSerProLeuPro----- 56  
Qy 666 CCCCAGGAATCTGGCCAGCTGCTCCACCATCGTCCACCAAGTCTC----- 624  
Db 57 ProProSerGlnProProProProLeuProLeuSerPro-LeuLysLysLysAspAsnIle 76  
Qy 623 -----CCGATGATGCTTTCATATCTTGAGAAAAAGCAGCTTCGAGTCCATGC 577  
Db 76 eGluLeuTrpLeuProValIlePhePheMetValLeuValAlaLeuValGlyMetGlyLe 96  
Qy 576 GGTGCTCCAGAACTCTCTGACAGCTGCTCAAAACAGGAGCC----- 537  
Db 96 uGlyMetTyrGlnLeuPheHisLeuGlnLysGluLeuAlaGluLeuArgGluPheThrAs 116  
Qy 536 -----GTAGCCCACTTATCATTTATAAACCTCCGGCAGA 502  
Db 116 nHisSerLeuArgValSerSerPheGluLysGlnIleAlaAsnProSerThrProSerGl 136  
Qy 501 ATTCTTGAAGAGTGCCTCGGTAGTCTCAT-----GGTCCGAGAGAGCGGTGGAACCTCAT 448  
Db 136 uThrLysLysProArgSerValAlaHisLeuThrGlyAsnProArgSerArgSerIlePr 156  
Qy 447 AATAAGACACCAACAGATCTCTGGGTTGGCGACCATATAGATGACCTTGGAGTCTCCAT 388  
Db 156 oLeuGluTrpGluAspThrTyrGlyThrAla-----LeuIle 168  
Qy 387 TGTGAGGTTCAGAGGCGCAGAAAGCGGTAGGCGAGTGGCTCTTGTATAGGCGGG----- 333  
Db 168 eSerGlyValLysTyrLysLys-----GlyGlyLeuValIleAsnGluAlaGlyLeuTy 186  
Qy 332 -----AGAGGTTCAG 324  
Db 186 rPheValTyrSerLysValTyrPheArgGlyGln 197  
RESULT 55  
US-10-861-934-26  
; Sequence 26, Application US/10861934  
; Publication No. US20060089491A2  
; GENERAL INFORMATION:  
; APPLICANT: Nagata, Shigikazu

666 CCCCAGGAATCTGGCCAGCTGCTCCACCATCGTCACCAAGTC-----

Qy 717 -----GCTGGTGCGAGTCTCGTCAGGGCTTCAGCTGGG---CCTTGT----- 676



```

Db      118 ValGlyHisArgGlySerProProArgLeuProGlyValProValIleSer 137
Qy      675 -----CACAGACACCCCGAGGATCTGGCCAGCTGCTCCACCATCTCCACAGGTCCC 622
Db      138 LeuGlyHisSerArgProProGlyVal-----AlaMetProThrThrGluLeuGly-Th 155
Qy      621 GATGCATGCTTCATACCTTGAGAAAAAGCAGCTTCGAGTCATCGGTGGTCCCGAGAACT 562
Db      155 rLeuArg-----ProProLeu 160
Qy      561 CCTGCACGTGCTCAAAACCCAGGA---GCCGTAGCCAGCTTATCATTTATAAACCTCCGGC 505
Db      160 uLeuGlnLeuSerThrLeuGlyThrAlaProProThrLeuAlaLeuHisTyrHisProHi 180
Qy      504 AGAATCTTTGAAAGTGCTCGGTAGCTCATGTGTCGCGAGAGCGGTGGAACTGATTAAT 445
Db      180 sProPheLeuAsn-----SerValTyrIleGlyProAlaGlyProPheSer---IlePh 197
Qy      444 AAGACACCAGATCCTTGGGGTTGCGAGCCATATAGATGACCTTGGAGTCTCCATTGT 385
Db      197 eProSerSerArgLeuLysArgArgProSerHisCysGlu-----LeuAspLeuAl 214
Qy      384 GCAG---GTCAGAGGCGAGAAAGCGGTAGGCGAGTGGCTCTTGAT-----GAGGCGGG 334
Db      214 aGluGlyHisGlnProGlnLysValAlaArgValPheThrAsnSerArgGluArgTr 234
Qy      333 GAGAGGTGAGTTCCTTGATGATGTCCAGGCCGCTGTGGGTACTC----- 288
Db      234 pArgGlnGlnAsnValAsnGlyAlaPheAlaGluLeuArgLysLeuLeuProThrHisPr 254
Qy      287 -CAGACCGGAGAGTCTGCTCGATGTTTCATCAAGCCGATCTCATCGGGTTCAGGCCCT 229
Db      254 oProAspArgLysLeu-----SerLysAsnGluValLeuArgLeuAlaMetLysTyrIl 272
Qy      228 GCCTCACCAAGTAGACCACTCTCGACCAAGCTGGTCCGAGCTTGGGTAGGTGACGA 169
Db      272 eGlyPheLeuVal---ArgLeuLeuArgAspGlnAlaAlaLeuAlaGlyProTh 291
Qy      168 TCCACACGTCTCGGCGC-----CACCGGAAGTTGGCGATCT-CCTCC 125
Db      291 rPro-----ProGlyProAspGlyArgProValHisArgValProAspArgGlyProAr 309
Qy      124 ATCTTCCCGCGCAGAGGCGGCGAGCCGCA 94
Db      309 gArgGlySerGlyArgAlaGluAlaAla 319

```

## RESULT 57

```

US-10-505-928-284
; Sequence 284, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 284
; LENGTH: 620
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-284

```

```

Alignment Scores:
Pred. No.:      2.75      Length:      620
Score:          88.00     Matches:     125
Percent Similarity: 32.8%   Conservative: 51
Best Local Similarity: 23.3% Mismatches: 192
Query Match:      2.0%     Indels:     169

```

```

DB:      6      Gaps:      27
US-10-768-158-1 (1-2419) x US-10-505-928-284 (1-620)
Qy      1702 CTGTGGAATTCAGCTCAAAATTCGAGCGCAACTCGGTCTCTTTGTGTGCGCCGACTCTCCACC 1643
Db      55 ValTrpGlyAlaGluGlyIleProAlaProThrCysTrpIleGlyThrAsp----- 71
Qy      1642 CACTGCATGCAAAATGCTCCTGACTGCATTAGACTCCCACTGCGCAGCTCACCAGGAGGAGA 1583
Db      72 -----ProGlySerProSerArg 77
Qy      1582 CTTCCTCTCAGCACATGCTCGCGGGCT-----CCTCTGCTGTCATGGGCC 1541
Db      78 -----AlaHisGlnProGlnAlaSerAspAlaAsnArgGluProVal-AlaGluAr 94
Qy      1540 TTTGAGCCGCCGCTCGAAGTCTCTGGCAGAGGACGTCCATTCTCAAGACCTAGTGA 1481
Db      94 gSerGluProAlaLeuSerGlyLeuProProAlaThrMetGlySerGlyAspLeuLeu-- 113
Qy      1480 CAATTTCTGCTCATTTTCAGGAAGGTTCTGAAGAAGTGAACCTCCGTCAGAATAACAAGTG 1421
Db      114 -----LeuSerGlyGluSerGlnValGluLysThrLysLeuSerSerSe 128
Qy      1420 TCCAGATTCCCCCAACAGACCCCAACTC-----ATCCTCTGTTCTTCCACCACTA 1373
Db      128 rGluGluPheProGlnThrLeuSerLeuProArgThrThrLysCysSerGlyHisAs 148
Qy      1372 TCTCAGAACCAAGTAGAAGACAGAACTGCTGT-----GCCTTAGATAATAAACA 1325
Db      148 pAlaAspThrGluAspAspProSerLeuAlaAspLeuProGlnAlaLeuAspLeuSerG1 168
Qy      1324 TACAAAAATAAGAGTTAACCTGAGATCTTTTACCATCTCGGTGTGGTCTGCTCTGA 1265
Db      168 nGlnProHisSerSer-----GlyLeuSerCysLeuSerGlnTrpLysSerValLe 185
Qy      1264 TTCCCTCTGGAAATGAATCTTTTATTTGTTTACTGACATTTATGTAGATTTTCCAGTGAAA 1205
Db      185 uSerPro----- 187
Qy      1204 AGCTCTATAAAATACAAATAATACGCGGTGTAAGAGGAGCAGACATTCTAGTTGCATATATTA 1145
Db      188 -----GlySerAlaAlaGlnProSerSerCys-SerIle- 198
Qy      1144 CAGGCTTTATCTTACGTCAGGCCATTTGAACTGCAATGTGGAGACTGTTTGTAAATCA 1085
Db      199 -----SerAlaSerSerThrGlySerSerLeuGln-----G 209
Qy      1084 GACATGGAGAGGCTGCACGTTCTAAAGCGCAGACAGCTTCTCGTTGGGAATCATCACA 1025
Db      209 lyHisGlnGluArgAlaGluProArgGlySerLeuAlaLysVal---SerSer-Ser 227
Qy      1024 CTCCCTCCGCTCACGCGCTCTTCCCTTCCCGCTGTTTCA-----CACGCTGCT 974
Db      228 LeuGluProValProGlnGluProSerSerValValGlyLeuGlyProArgProGln 247
Qy      973 TCCAGAGTTTGTCCAGCAAGGAATAAATGAATGATCATCAGGACTTTGGCTAGTAGACTG 914
Db      248 TrpSerProGlnProValPheSerGlyGlyAspAlaSerGlyLeuGlyArgArgLeu 267
Qy      913 TCTGGGTATTGTGAGCATGCGGTTGTTGTTTCTGTTATTATTAATAAATAAAGTCAACGTG 854
Db      268 SerPheGlnAlaGluTyrTrpAlaCys-Val----- 277
Qy      853 AGGTACACACTTTTCCCATCTTCTGTTTATACACCAAGTCAAACTTCTCATTTCA-----TG 800
Db      278 -----LeuProAspSerLeu---ProProSerProAspArgHisSerProLeuTr 293
Qy      799 GAGACGGTGAAGATGTCCTTCC-----ACAGCCCAACT 767
Db      293 pAsn-----ProAsnLysGluTyrGluAspLeuLeuAspTyrThrTyrProLe 309
Qy      766 CTTCCCGCGCCCGGCGGCGGCTCAGCGGTGCGACACTGGTCCACAGCT----- 715

```

```
Db 309 uArgProGlyProGlnLeu---ProLysHisLeuAspSerArgValProAlaAspProVa 328
Qy 714 -----GGTGGCAGTGTCCGTGAGGCTTCCAGCTGGGCTTGTGCAC-----AG 671
Db 328 lLeuGlnAspSerGlyValAspLeuAspSerPheSerValSerProAlaSerThrLeuLy 348
Qy 670 GACACCCCGCAGGATCGCCAGCTGTCTCCACCATCGTCACCAAGTCCCGATGATGTCT 611
Db 348 sSerProThrAsnValSerProAsnCysProAla----- 360
Qy 610 TCATACTTGAAAGAACAGCAGTCCAGTCCATGCGGTGCTCCAGAACTCCTGCACGTGC 551
Db 361 -----GluAlaThrAlaLeuProPheSerGlyProArgGluProSerLeuLy 376
Qy 550 TCAAAACGAGGAGCGGTAGCCAGCTTATCATATCAATAAACCTCCGCGAGAAATCTTGAAG 491
Db 376 sGlnTrpProSerArgValProGlnLysGln-----GlyGlyMetGlyLeu----- 391
Qy 490 GTGCTCGGTAGCTCATGCTCCGAGAGAGCGGTGGAACTGATATTAAGACACCAACAGA 431
Db 392 -----AlaSerTrpSerGlnLeuAla-----SerThrPro----- 401
Qy 430 TCCTTGGGTTCGAGCATATAGATGACCTTGGAGTCTCCATTTGTGGAGTCAAGAGGC 371
Db 402 -----ArgAlaProGlySerArgAspAlaArgTrpGluAArgAr 414
Qy 370 AGAAAGCGGTAGCGAGTGGCTTTGATGAGCGGGAGAGTCAAGTCTCTTGTATGATG 311
Db 414 gGluProAlaLeuArgGly-AlaLysAspArgLeuThrIleGlyLysHisLeuAsp-Met 433
Qy 310 TCCAGGCCCGGTGTG-----GGTACTCCAGGACCGGAGTGTCTGCTCG 266
Db 434 Gly-SerProGlnLeuArgThrArgAspArgGlyTrpProSerProArgProGluAArgL 453
Qy 265 ATGTTTCATCAAGCGATCTCATCGGGTTCAGCGCCT 229
Db 453 uLysArgThrSerGlnSerAlaArg-----ArgPro 463
```

## RESULT 58

```
US-10-511-937-2446
; Sequence 2446, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 50661200104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2446
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2446
```

Alignment Scores:

Pred. No.: 2.41 Length: 345

```
Score: 87.50 Matches: 64
Percent Similarity: 30.8% Conservative: 22
Best Local Similarity: 22.9% Mismatches: 86
Query Match: 2.0% Indels: 107
DB: 6 Gaps: 15
```

US-10-768-158-1 (1-2419) x US-10-511-937-2446 (1-345)

```
Qy 762 CCAGCCCGCAGCGGCTCAGCGTTCAGCACTGGTCCACAGCTGGT-----GGC 709
Db 50 ProAlaAlaArgProGlyProArg-----ProProAlaGlyGluLeuGly 64
Qy 708 AGTGCTCCGTTCAGGGCTTCAGCGTGGGCTTGTTCACAGGACA-----CCCCCAGGA 658
Db 65 Ser-----lIleGlyAspHisGluArgAlaAlaPheSerProTyr 78
Qy 657 ATCTGGCCAGCTGCTCCACCATCGTCCAGAGTCCCGATGCATGTCTTTCATCTAGAA 598
Db 79 LeuGluProLeuGlyAlaProGlnAlaProAlaProAlaThrAlaThrAspThrPheGlu 98
Qy 597 AAAGCAGCTTCGAGTCCATCGGTCTCCAGAACTCCTGCACGT----- 553
Db 99 AlaAlaProProAlaProAlaProAlaProAlaSerSerGlyGlnHisAspPheLeu 118
Qy 552 -----GCTCAACACGAGGAGCGGTAGCCAGCTTATCATTTCA 517
Db 119 SerAspLeuPheSerAspAspTyrGlyGlyAsnCysLysLysProAlaGluTyrGly 138
Qy 516 TAAACCTCC---GGCAGAAATTTTGAAGGTGCTCGGTAGCTCATCGTCCGAGAGAGC 460
Db 139 TyrValSerLeuGlyArg-----LeuGlyAlaAlaLysGlyAlaLeuHis 153
Qy 459 GTTGGAACTGAT-----AATAAGACACCACTGCT----- 427
Db 154 ProGlyCysPheAlaProLeuHisProProProProProProProAlaGluLeu 173
Qy 426 -----TGGGTTTCGAGCCATATAGATGACCTTGGAGTCTCCATTTGGAGGTCAG 376
Db 174 LysAlaGluProGlyPheGluPro-----AlaAspCys---LysArg 186
Qy 375 AGGCGAGAAAGCGGTAGGCGAGTGGCTCTTGTATGAGCGGGGAGAGGTTCAGTTCTTGA 316
Db 187 LysGluGluAlaGlyAlaProGlyGly----- 196
Qy 315 TGATGTCCAGCCCGCTGTGGTACTCCA----- 286
Db 197 -----AlaGlyMetAlaAlaGlyPheProTyrAlaLeuArgAlaTyrLeuGlyTyrGln 214
Qy 285 -----GGACCGGAGCTGCTCGTCGATGTTTCATCAAGCCGATCTCATCGGGT 238
Db 215 AlaValProSerGlySerSerGlySerLeuSerThrSerSerSerSer----- 231
Qy 237 CAGCGCCTTGCTCACCAGTAGACACCTCTCTCGACAAAGCTGGTCCCGAGTGGGGT 178
Db 232 ---ProGlyThrProSer----- 237
Qy 177 AGGTGACGATCCACACGCTGCTGGCGCCGACCGGGAAGTTGGCGATCTCTCCATCTTCC 118
Db 238 -----ProAlaAspAlaLysAlaProProThrAla 247
Qy 117 CGCGGAGAAAGGCGGAGCGGCGGAGCGGCGCATGGAACTCGAAGTACTTGTCTCGAACT 61
Db 248 CysTyrAlaGlyAlaAlaProAlaProSerGlnValLysSerLysAlaLysLysThr 266
```

## RESULT 59

```
US-11-101-316-16
; Sequence 16, Application US/11101316
; Publication No. US20060099657A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
```



```
Score: 87.50 Matches: 79
Percent Similarity: 35.9% Conservative: 47
Best Local Similarity: 22.5% Mismatches: 126
Query Match: 2.0% Indels: 99
DB: 6 Gaps: 16

US-10-768-158-1 (1-2419) x US-10-196-749-440 (1-1115)
Qy 1032 TCATCAGACTCCCTCCGCTCAGCGGCTCTTCCCTCCCGCTGTTTCACAGCTGCTT 973
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 365 SerSerGlnArgLeuArgLeuSerArgAlaLeuValLeu-----SerMetGly 382
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 972 CCAGAGTTTCCAGCAGGAATAAATCAATGATACATACAGGACTTTGGCTAGTAGACTGT 913
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 383 ProGlu-----AspGluGlyValTyGlnCysMetAlaGluAenGluValGlySerAla 400
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 912 CTGGGTATTGTGAGCATGACGAGTTGTTGTTCTGTTATTATAAATAAAGTCAAACTGTA 853
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 401 HisAlaValGlnLeuArg-----ProGluGlyValTyGlnCysMetAlaGluAenGluValGlySerAla 400
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 852 GGTCACTTCCCATCTCTGTTTATACACCAAGTCAAACTTCTCATTCATGAGACGG 793
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 408 -----ThrSerArgProSerIleThrProArgLeuTrpGlnAsp 420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 792 TGAAGA---TGCTCTCCACAGCCCACTCTCCCGGCCACCGGCGGCGCTCAGCGT 736
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 421 AlaGluLeuAlaThrGlyThrProValSerProSerLysLeuGlyAsnProGluGln 440
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 735 TGCAGCACTGGTCCACAGCTGGTGGCAGTGCTCCGTCAGGCTTCCAGCTGGGCTGT 676
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 441 MetLeuArgGlyGlnProAla-----LeuPro----- 449
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 675 CACAGGACACCCCGCAGGAATCTGGCCAGCTGCTCCACCATCGTCACAGGTCCCGATGCA 616
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 450 ---ArgProProThrSerValGlyProAlaSer---ProLysCysProGly----- 464
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 615 TGTCTTCATCTTGAGAAAGACAGCTTCGAGTCCATCGCTGCTCC-----AGA 565
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 465 -----GluLysGlyGlnGlyAlaProAlaGluAlaProIleLeuSer 479
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 564 ACTCCTGCAGCTGCTCAACACGAGCGGTAGCCAGCTTATCATATCAATAAACCCTCCGCG 505
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 480 SerProArgThrSerLysThrAspSer-TyrGlu-----LeuValTrpArgProArgHi 497
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 504 AGAATCTTGAAGGTGCTCGGTAGCTCATGCTCCGACGAGAGCGGTGGAACTG----- 450
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 497 sGlu-----GlySerGlyArgAlaProIleLeuTy 508
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 449 -ATAATAGACACCAACAG-----ATCCTTGGGCT 421
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 508 rValValLysHisArgLysGlnValThrAsnSerSerAspTrpThrIleSerGlyI 528
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 420 TGCAGCCATATAGATGACCTTGGAGTCTCCATTGTGGAGTTCAGAGGCGAGAAGCGGT 361
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 528 eProAlaAsnGlnHisArgLeuThrLeuThrArgLeuAspProGlySerLeuTyGluVa 548
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 360 AGGCGAGTGGCTCTTGATGAGCGGCGGAGAGGTTCAGTTCCCTTGATGATGCCAGGCCG 301
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 548 lGluMetAlaAlaTyAsnCysAlaGlyGluGln-ThrAlaMetValThr-PheArg 567
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 300 GCTGTGGGTACTCCAGGACCGGGA-----GCTGCTCGTCGATGTTTCATCAAGCGATCT 247
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 568 ThrGlyArgArgProLysProGluLeuMetAlaSerLysGluGlnIleGlnArgAsp 587
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 246 CATCGGGGTACGCCCTGGGTCCACCAAGTAGACCACTCTTCGACAGCAAGCTGGTGGCG 187
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 588 AspProGlyAlaSerProGlnSerSerGlnPro----- 599
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 186 ACTTGGGTAGTGCAGATCCACAGCTCGCTGGGCGCGCACCGGGAAGTTGGGATCTCCT 127
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 600 AspHisGlyArg-----LeuSerPro 606
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 126 CCATCTTCCCGCGGAGAGGCGGCGAGCCGACCGCATGCCATGGAACTCGAAGTACTTGCTCT 67
```

```
Db 607 ProGluAlaProAspArgProThrIleSerThr-AlaSerGluThrSerValTyValTh 626
Qy 66 CGAACTCCCCCGGGTCTCTGGGGT 42
Db :::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 626 rTrpIleProArgGlyAsnGlyGly 634
RESULT 61
US-11-302-678-11
; Sequence 11, Application US/11302678
; Publication No. US2006008881A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: Venkateswarlu, Karicheti
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 1435, 559, 34021, 44099, 25278,
; TITLE OF INVENTION: 641, 260, 55089, 21407, 42032, 46656, 62553, 302, 323,
; TITLE OF INVENTION: 12303, 985, 13237, 13601, 18926, 318, 2058 OR 6351 MOLECULES.
; FILE REFERENCE: MPI02-012P1RNM OMNI
; CURRENT APPLICATION NUMBER: US/11/302,678
; CURRENT FILING DATE: 2005-12-14
; PRIOR APPLICATION NUMBER: US/10/345,680
; PRIOR FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US 60/349,511
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/360,500
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/365,041
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/374,063
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/403,468
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US 60/414,262
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: US 60/419,986
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US 60/423,809
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 60/429,797
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-302-678-11

Alignment Scores:
Pred. No.: 2.97 Length: 459
Score: 87.00 Matches: 42
Percent Similarity: 31.8% Conservative: 12
Best Local Similarity: 24.7% Mismatches: 60
Query Match: 2.0% Indels: 56
DB: 7 Gaps: 8

US-10-768-158-1 (1-2419) x US-11-302-678-11 (1-459)
Qy 1002 TCCCTTCCCGCTGTTTTCACACGCTGCTTCCAGAGTTTGTCCAGCAAGGAATAAATGAA 943
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 325 SerLeuIleProThrIleIleAsnLeuAlaThrAlaLeuThrSerValGlyValGlySer 344
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 942 TGCATACAGGACTTTTGGCTAGTAGACTGTCTGGGTATTGTGAGCATGCGAGTTGTTGTT 883
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 345 PheLeuCysAsp---Trp-----IleLeu 351
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 882 TCTGTTATTATAAAGTCAACGTCAGGTGCACACTTTCCCATCTTCTGTTTATACA 823
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 352 LeuThrPheMetAsnLysAsnLysValTySerHisLysLysPheAspLysValCysThr 371
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 822 CCAAGTCAAACCTTCTCATTCATGAGAGCGGTGAAGATGTCTCTCCACA----- 775
```



QY	184	AGTCCGGCACCAAGCTTGCTGCGAGGAGGTGGTTACTATTGGTAGCACCAGGCCTGACCCCG	243
DB	558	ProLeuProGlyMetLeu-----	569
QY	244	ATGAGTCGGCTTGATGAACATCGACGAGCAGCTCCCGGTCTCTGGAGTACCACACGCCG	303
DB	570	-----ProProProProProProProProProProGlyyl	580
QY	304	GCTTGGACATCATCAAGAACTGACCTCTCCCGGCTCATCAAGAACCACTGCCCCTACC	363
DB	580	ypro-----ProProProProGlyyProPro-----Pr	589
QY	364	GTTTTCTGCTCTGACTCCA-	385
DB	589	oLeuGlyAlaIleMetProProProGlyAlaProMetGlyLeuAlaLeuLysLysLysLeSe	609
QY	386	-----CAATGG-----AG	393
DB	609	rileProGlnProThrAenAlaLeuLysSerPheAsnTrpSerLysLeuProGluAsnily	629
QY	394	ACTCCAAGGTCTATATATGGCTCGCAACCCCCAAGGATCTGTTGGTGTCTTATTACGT	453
DB	629	sLeuGluglyThrValTrpThrGluIleAspAspThrLysValPheLysIleLeuAspLe	649
QY	454	TCCACCGCTCTCTCGGGACCATGACTACCGAGGCACCTTCAAGAATACTCCCGAGGT	513
DB	649	uGluAspLeuGlu-ArgthrPheSerAlaTyrglNargInglndlnasphe-----p	666
QY	514	TTATGAAT-	525
DB	666	hevalasnSerasnSerlysGlnlysgluAlaAseAlaIleaspsphrLeuSerSerL	686
QY	526	AGCTGGGTACGGCTCTCTGGTTTGAGCACGTGCAGGAGTTC-----TGGGAGC	573
DB	686	ysLeu-----LysValLysGluLeuSerValIleAspGlyArgA	699
QY	574	ACCGATGGACTCGNACTGCTTTT-----CTCAAGTATGAACATGCATC	621
DB	699	rgAlaGlnAsnCysAsnIleLeuLeuSerArgLeuLysLeuSerAsnAspGluIleLysA	719
QY	622	GGGACCTGGTGCAGC--	651
DB	719	rgAlaIleLeuThrMetAspGluGlnGluAspLeuProlysnspMetLeuGluGlnLeu	739
QY	652	CCAGATCTCTGGGGTGTCTGTGACAAGGCCAGCTGGAAGCCTGACGGAGCACTGCC	711
DB	739	eUlysPheValpro-----GluLysSerAspileAspLeuLeuGluGluHisLysH	756
QY	712	ACCAGCTGGTGAGCAGCTGCTGCAACGCTGAGGCCCTGCCCGTGGCGGGGAAGTTG	771
DB	756	iagGluLeu---AspargMetalaLyAspArgPheLeuPheGluMetSerArglleA	775
QY	772	GGCTGTGGAAGGAC---ATCTTCAAGTGTCCATGAATGAGAAGTTT	815
DB	775	snHistyrGlnGlnArgLeuGlnSerLeuTyrrPheLysLysLysPhe	790
RESULT 64			
US-11-301-554-1863			
; Sequence 1863, Application US/11301554			
; Publication No. US20060088527A1			
; GENERAL INFORMATION:			
; APPLICANT: Henderson, Robert A.			
; APPLICANT: Wang, Tongcong			
; APPLICANT: Watanabe, Yoshihiro			
; APPLICANT: Kalos, Michael D.			
; APPLICANT: Sleath, Paul R.			
; APPLICANT: Johnson, Jeffrey C.			
; APPLICANT: Retter, Marc W.			
; APPLICANT: Durham, Margarita			
; APPLICANT: Carter, Darrick			
; APPLICANT: Fanger, Gary R.			
; APPLICANT: Vedvick, Thomas S.			

QY 345 TGATGAGCGGGGAGAGTCACTTCTTGATGATGCCAGGCCGGCTGTGGTACTCCA 286  
Db |||||  
216 -----GlyGlyGlu-:: ||||  
QY 285 GGACCGGAGCTGCTCGTGTTCATCAAGCCGATCTCAGGGTCAAGCCCTGGC 226  
Db |||||  
224 AspAspGlyAlaAlaCysArgArgAsnAlaGlyGlnGlyArgArgGlySerGlyGlyAla 243  
QY 225 TCACCAAGTAGACCCTCTGCGCAAGCTGGTCCGCGACTTGGGTAGGTGACGATCC 166  
Db |||||  
244 ArgGlyAlaAlaGlu-ArgArgAlaGlyArg-:: |||||  
QY 165 ACACGTCTGGCGCGCAGCGGAAGTGGCGCATCTCTCATCTTCCGCGGCGAGAAG 106  
Db |||||  
256 nHisProLeuGlyProHisArgArg--GlyAlaGlnArgAlaAlaGluArgAlaHisPro 275  
QY 105 GCGGCGCGCGCAGCATGGAAGTCTCGAAGTACTTCTCTCGAATCCCGCGGGTCTGG 46  
Db |||||  
276 AlaAlaAlaValArgValGlyProArgGlnGlyAlaGluProArgGlyHisAspProGly 295  
QY 45 GGGTCTCGGCTCGCTCTCCG 25  
Db |||||  
296 GlyProArgGlnArgAlaPro 302  
RESULT 65  
US-10-511-937-2549  
; Sequence 2549, Application US/10511937  
; Publication No. US2006008836A1  
; GENERAL INFORMATION:  
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.  
; APPLICANT: Wohlgemuth, Jay  
; APPLICANT: Fry, Kirk  
; APPLICANT: Woodward, Robert  
; APPLICANT: Ly, Ngoc  
; APPLICANT: Prentice, James  
; APPLICANT: Morris, Macdonald  
; APPLICANT: Rosenberg, Steven  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING  
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION  
; FILE REFERENCE: 506612000104  
; CURRENT FILING DATE: US/10/511,937  
; CURRENT FILING DATE: 2004-10-19  
; PRIOR APPLICATION NUMBER: PCT/US2003/012946  
; PRIOR FILING DATE: 2003-04-24  
; PRIOR APPLICATION NUMBER: US 10/131,831  
; PRIOR FILING DATE: 2002-04-24  
; PRIOR APPLICATION NUMBER: US 10/325,899  
; PRIOR FILING DATE: 2002-12-20  
; NUMBER OF SEQ ID NOS: 3117  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2549  
; LENGTH: 394  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-511-937-2549  
Alignment Scores:  
Pred. No.: 3.39 Length: 394  
Score: 86.00 Matches: 59  
Percent Similarity: 32.1% Conservative: 22  
Best Local Similarity: 23.4% Mismatches: 89  
Query Match: 2.0% Indels: 82  
DB: 6 Gaps: 11  
US-10-768-158-1 (1-2419) x US-10-511-937-2549 (1-394)  
QY 1 GCGACGGCGCGCGCGCATGCGGAGAGCGAGCGCCGACACCCCGACCGCGGGG 60  
Db |||||  
186 AlaHisAlaAlaAlaGlyProGlyGluValLeuAlaLysGluProAlaProArg--- 204  
QY 61 AGTTCGAGACGACGACTTCTCGAGTCCATGGCGTGGCTGCGCGCTTCTGCGCGGGA 120  
|||

Db 205 -----AlaPro 206  
QY 121 AGATGGAGGAGATCGCCAACTTCCGGTGCAGCGCCAGCGAGTGTGGATCGTCACTACC 180  
|||  
207 ArgAlaThrArgAlaProValAla-----SerProAlaAlaLeuGlySerThrAlaThr 224  
QY 181 CCAAGTCCG---GCACGAGCT-----TGCTGCGAGGAG 210  
|||||  
225 AlaSerProAlaAlaProAlaArgArgGlyLeuGluCysSerGluCysLysLysLeu 244  
QY 211 TGGTCTACTTGGTGAGCCAGCGCGTGCACCGATGAGATCGGCTTGATGAACATCGACG 270  
::: |||||  
245 PheSerThr----- 247  
QY 271 AGCAGCTCCCGTCTCGAGTACCACAGCGCGCTGGACATCATCAAGAACTGACCT 330  
|||  
248 -----GluThrSerLeuGlnValHisArgArgIleHisThrGlyGluArgProTyrPro 265  
QY 331 CTCCCGCGCTCATCAAGAGCCACCTGCCCTACCGCTTCTGCCCTCTGACCTCCCAATG 390  
|||  
266 CysPro-----AspCysGlyLysAlaPheArgGlnSerThr--HisLeuL 280  
QY 391 GAGACTCCAAGTGCATCTATATGCTCGCAACCCCAAGGATCTGCTGCTGCTTATTATC 450  
|||  
280 ysAspHisArgArgLeuHisThrGlyGluArgProPheAlaCysGluValCysGlyLysA 300  
QY 451 AGTTCACCGCTCTCTCGGACCATGAGCTACCGAGCACCTTTCAAGATTTCTGCCGGA 510  
|||  
300 laPheAlaIleSerMetArgLeuAlaGluHis-----ArgA 312  
QY 511 GGTATTGAATGAATAAGCTGGGTAC-----GGCTCTGGTTTGGACAG 555  
|||  
312 rgIleHisThrGlyGluArgProTyrSerCysProAspCysGlyLysSerTyrArgSerP 332  
QY 556 TGCAGGAGTTCTGGGAGCACCGCTGAGTCAAGCTGCTTTCTCAAGTATGAAGACA 615  
Db 332 heSerAsnLeuTrpLysHisArgLys-----ThrHisGlnGlnG 345  
QY 616 TGCATCGGGACCTGCTGACGATGCTGGAGCAGCTGGCCAGA-----TTCTCTGGGGG 666  
|||  
345 lnHisGlnAlaAlaValArg-----GlnGlnLeuAlaGluAlaAlaValGlyL 363  
QY 667 TGTCTGTGACAAAGCCCGACCTGGAAGCCCTG 698  
Db 363 euAlaValMetGluThrAlaValGluAlaLeu 373  
RESULT 66  
US-11-246-999-48  
; Sequence 48, Application US/11246999  
; Publication No. US20060099622A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 12 Human Secreted Proteins  
; FILE REFERENCE: PF489P2  
; CURRENT APPLICATION NUMBER: US/11/246,999  
; CURRENT FILING DATE: 2005-10-11  
; PRIOR APPLICATION NUMBER: US/09/984,130  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/243,792  
; PRIOR FILING DATE: 2000-10-30  
; PRIOR APPLICATION NUMBER: 09/836,353  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/198,407  
; PRIOR FILING DATE: 2000-04-19  
; PRIOR APPLICATION NUMBER: PCT/US99/25031  
; PRIOR FILING DATE: 1999-10-27  
; PRIOR APPLICATION NUMBER: 60/105,971  
; PRIOR FILING DATE: 1998-10-28  
; NUMBER OF SEQ ID NOS: 149  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 48  
; LENGTH: 995  
; TYPE: PRT

; ORGANISM: Homo sapiens  
US-11-246-999-48

## Alignment Scores:

Pred. No.: 4.86 Length: 995  
Score: 86.00 Matches: 89  
Percent Similarity: 33.0% Conservative: 48  
Best Local Similarity: 21.4% Mismatches: 135  
Query Match: 2.0% Indels: 143  
DB: 7 Gaps: 19

US-10-768-158-1 (1-2419) x US-11-246-999-48 (1-995)

```
QY 2298 CCATACAGTACTAAATAGAAAAATATAATGAAATTTACAGAAAGCAGCCCTCCCTCA 2239
|||
DB 646 ProThrLysAlaProArgLysThrProThrLysPheProThrLysValProProThr 665
|||
QY 2238 CAGAAACACAGGCAAGGTGCTGCCGAGCCACC----- 2206
|||
DB 666 ProThrLysPheProThrLysValProProThrProThrGlnPheProThrLysValPro 685
|||
QY 2205 -----ATCCGATGCCAGCTGCCACTGCTGCCAGCAGCTTC----- 2170
|||
DB 686 ProLleProThrLysValProSerLysValProProThrProThrGlnPheProThrLys 705
|||
QY 2169 -----TCCCATTAACAACAAAACACAAAGCTTCTCTACACC----- 2130
|||
DB 706 ValProProThrProThrLysValSerThrLysValLeuSerThrProThrLysAlaHi 725
|||
QY 2129 -----GTTTCAATACAGCACCAAGACGAGACCCCTTCCACAGCAGCAGCCCTCCT 2077
|||
DB 725 sThrLysValSerProThrSer-----ThrLysLeuProSerLysAlaProSer----- 741
|||
QY 2076 GACCACCACTTTGGGTGCTTTCCAGCCACATAAATAGACAGAACCCCGTCCCAACGGG 2017
|||
DB 742 -ThrGlnThrMetValProThrLysValHis-----ProThrProThrLys 756
|||
QY 2016 GTTGCTGGCTACTGCGCCCGCAGCTCTCCATCGCTTCAGCAATGTGGTGAGAAATTA 1957
|||
DB 756 sLeuProThrLysValProProLleThrThrLysValSerAsn----- 770
|||
QY 1956 GGCAGTGAAGCAATAAAGACCTCACCTTTACTCTCCCTCAGATACACGGAGGATGAGGTG 1897
|||
DB 771 -LysValLeuLeuThrSerProGluLeu----- 779
|||
QY 1896 ACAGACAGCAGCGCGCGTCCCAACCTGATCCGAGTGAAACAGGGCTACACTCGCAAAA 1837
|||
DB 780 -----ProThrLysValProProThrProThrLysLeuProThr--As 793
|||
QY 1836 TGGTCTCCACGGCTGAGCGCTCAGGTGCACAGAACAACTGAATATAAGACCCGG 1777
|||
DB 793 nAlaProProThr----- 797
|||
QY 1776 GGATTTTAGAATTTACTATTAAAG-----ATCTGCTTGCAGCATTAGTAGGCTA 1726
|||
DB 798 -----SerLleLeuLeuSerProThrProLleLysLeuProThrLysIle-SerL 814
|||
QY 1725 TCACCTTCAGGAAC----- 1712
|||
DB 814 euThrLeuThrSerValProLleLysAsnGlnLeuThrSerAlaLysLeuLeuThrThr 834
|||
QY 1711 -----AAAGTAAGGTGTGGAATTCAGTCAATTCGAGCGCAACTCGTCTTTGGTG 1657
|||
DB 834 hrLeuProLleSerThrLysArgAlaThrLysLeuProSerThrSerThrSerValPro 854
|||
QY 1656 GCCGACTCTCC-----ACCCACTGCATCAAAATGCTCGT-----ACTGCAT 1615
|||
DB 854 eRAsnThrSerCysLleLeuThrHisValGlnProLysMetLeuProThrGluThrArgV 874
|||
QY 1614 TAGACTCCCACTGCCAGCCTCACCAGGAGACTTCCC-----TCAGCACATGCTCCGC 1561
|||
DB 874 alProAsnLysMetProProLysProThrArgLleProThrMetSerMetTyrlleThrL 894
|||
```

Alignment Scores:

Pred. No.:

5.81

Length:

1570

Score:

86.00

Matches:

75

```
QY 1560 GGGCTCCTCTGCAATTGGGCTTTGGAGCGCGCGCTGGAAGTCTCTGGCAGAGGACGTC 1501
|||
DB 894 yslsPro-----p 897
|||
QY 1500 CATTCTGAAGACCTAGTGAACAAATTTCTGTCATTTTCAGGAAGTCTTGAAGAAGTGAAC 1441
|||
DB 897 roLeuLysLysAsnSerAlaLysLysVal-----ThrAspL 909
|||
QY 1440 TCCGTGAGATAAACAAGTGTCCAGATTCCTCCCAACAGACCCCAACTCATCTCTGTCT 1381
|||
DB 909 ysArgProThrLysSerProLysThrThrLysProProLysProLysSerLysThrS 929
|||
QY 1380 CACCACCTATCTCAGAACCAAGTGAGAACAGAAATGGCTGTGCCTTAGATAATAAACT--- 1324
|||
DB 929 erValValAsnGln-ProLysLysLysGluThrLysThrGlyValAsnAsnLysThrLys 948
|||
QY 1323 -----ACAAAATAAAGAGTTAAACCTGAGATCTTCTACCATTCGGGTG----- 1279
|||
DB 949 AsnLeuProProLysAlaLysGluProLysLysLysLysThrLysValLysGln 968
|||
QY 1278 -----TGGCTCGCTCCTGATTCCTCCCTGGAAA 1252
|||
DB 969 ProValSerHisTyPheProProGlnLysProLysLys 981
|||
```

## RESULT 67

US-11-312-958-38  
; Sequence 38, Application US/11312958  
; Publication No. US20060100152A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Rosenfeld, Julie Beth  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING  
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 9949, 14230, 760, 62553,  
; TITLE OF INVENTION: 12216, 17719, 41897, 47174, 33408, 10002, 16209, 314, 636,  
; TITLE OF INVENTION: 72410, 33260, 619, 15985, 69112, 2158, 224, 615, 44373,  
; TITLE OF INVENTION: 95431, 22245, 2387, 16658, 55054, 16314, 1613, 1675, 9569 OR  
; TITLE OF INVENTION: 13424 MOLECULES  
; FILE REFERENCE: MP102-027P1ERNOMNIM  
; CURRENT APPLICATION NUMBER: US/11/312,958  
; PRIOR FILING DATE: 2005-12-20  
; PRIOR APPLICATION NUMBER: US/10/369,022  
; PRIOR FILING DATE: 2003-02-19  
; PRIOR APPLICATION NUMBER: US 60/360,495  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/370,121  
; PRIOR FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: US 60/373,010  
; PRIOR FILING DATE: 2002-04-16  
; PRIOR APPLICATION NUMBER: US 60/373,908  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: US 60/377,717  
; PRIOR FILING DATE: 2002-05-03  
; PRIOR APPLICATION NUMBER: US 60/379,949  
; PRIOR FILING DATE: 2002-05-13  
; PRIOR APPLICATION NUMBER: US 60/382,409  
; PRIOR FILING DATE: 2002-05-21  
; PRIOR APPLICATION NUMBER: US 60/385,280  
; PRIOR FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US 60/386,879  
; PRIOR FILING DATE: 2002-06-06  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 38  
; LENGTH: 1570  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-312-958-38

Alignment Scores:

Pred. No.:

5.81

Length:

1570

Score:

86.00

Matches:

75





```

Db      1620 AspThrTrpLeuThrThrAsnLeuPro-----AlaLeuThrValGlySerGlu 1635
Qy      451 AGTTCCACCGCTCTC-----TCCGGACCATGAGCTACCGAG 486
Db      1636 SerSerLeuAlaLeuAsgLeuValAsnGlyGlyAspArgCysArgGly---ArgValGlu 1654
Qy      487 GCACCTTTCAAGATTCTGCGGAGGTTTATGAATGATAAGCTGGGTACGGCTCCCTGGT 546
Db      1655 ValLeuTyrArgGlySerTrpGlyThrValCysAspAspSerTrpAspThrAsnAspAla 1674
Qy      547 TTGAGCAGCTGCAGAGTTCTGGGAGCACCAGCTGCACTCGAAGCTGC-----594
Db      1675 AsnValValCysArgGlnLeuGly---CysGlyTrpAlaMetSerAlaProGlyAsnAla 1693
Qy      595 -----TTTTTCAAGTATGAAGACATGATCGAATTCCTGG-----663
Db      1694 ArgPheGlyGlnGlySerGlyProileValLeuAspAspValArgCysSerGlyAsnGlu 1713
Qy      631 TGACGATGGTGAGCAGCTGCCAGATTCTGG-----663
Db      1714 SerTyrLeuTrpSerCysProHisGlyGlyTrpLeuThrHisAsnCysGlyHisGlu 1733
Qy      664 -----GGGTGCTCTGTGACAGGCGCCAGCTGGAAGCCCTGACGGAGCAGCTGCC 711
Db      1734 AspAlaGlyValIleCysSerAlaThrGlnIleAsn-----SerThrThr 1748
Qy      712 ACCAGCTGGTGG-----ACGAGTGTGCAACGCTGAGGCGCTGCCCGTGGCGGGGAA 765
Db      1749 ThrAspTrpTrpHisProThrThrThrThrAlaArgProSerSerAsnCysGlyGly 1768
Qy      766 GAGTTG-----GGCTGTGGAAGACATCTTCACGGTCT-----CCATGA 804
Db      1769 PheLeuPheTyrAlaSerGlyThrPheSerSerProSerTyrProAlaTyrTyrProAsn 1788
Qy      805 ATGAGAAGTTTACTTGGTGTATAACAGAAAGTGGAAAGTGTGACCTCA-----CGT 858
Db      1789 AsnAlaIysCysValTrpGluLe-----GluValAsnSerGlyTyrArg 1803
Qy      859 TTGACTTTTATTTATATAACAGAAACAACTGCATGCTCAATACCAACAGCAGTC 918
Db      1804 IleAsnLeuGlyPheSerAsnLeuLysLeuGluAlaHis---HisAsnCysSerPheAsp 1822
Qy      919 TAC 921
Db      1823 Tyr 1823

RESULT 69
; Sequence 383, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/19178
; CURRENT APPLICATION NUMBER: US/10/505,928
; PRIORITY FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 383
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-383

Alignment Scores:
Pred. No.: 5.19 Length: 720
Score: 85.00 Matches: 66
Percent Similarity: 29.2% Conservativity: 25
Best Local Similarity: 21.2% Mismatches: 119
Query Match: 1.9% Indels: 102

```

```

DB:      6      Gaps:      13
US-10-768-158-1 (1-2419) x US-10-505-928-383 (1-720)
Qy      822 CCAAGTCAAATCTTCATTTCATGAGACGAGTGAAGATGTCCTCCACAGCC-----772
Db      186 ProProHisSerAlaHisAlaLeuProArgGluSerCysThrAlaHisAlaAsnSerGln 205
Qy      771 -----CAACTCTTCCCGGC-----CCACGGGAGGCGCTCAGCGT 736
Db      206 AlaAlaThrGlnArgLysProGlyThrLysLeuLeuLeuProArgAlaAlaSerValArg 225
Qy      735 TGCAGCAGCTGCTCCACAGCTGGTGGCAGTCTCCGTCAGGCTCCAGCTGGGCTTGT 676
Db      226 GlyArgSerIleProGlyAlaAlaGluLysProLysGluIleProAlaSerProSer 245
Qy      675 CACAGGACACCCCGAG-----ATCTGGCCAGCTGTCTCCACCA 637
Db      246 ArgThrLysIleProAlaGluLysGluSerHisArgAspValLeuProAspLysPro---264
Qy      636 TCGTCACAGGT-----CCGATGTCATGCTTTCATCTTGCAGTGAAGAGCAGT 589
Db      265 ---AlaProGlyAlaValAsnValProAlaAlaGlySerHisLeuGlyGlnGlyLysArg 283
Qy      588 TCGAGTCCATGCGGTGCTCCAGAACTCTGCACGCTCTCAAAACAGGAGCGGTAGCCCA 529
Db      284 AlaIlePro-----ValProAsnLysLeuGlyLysLysThrLeuLysAlaPro 301
Qy      528 GCTTATCATTTAAACCTCCGGCAGAAATTTTGAAGGTGCTTCGTAGCTCATGTGTC 469
Db      302 GlySer-----ThrSerAsnLeuAlaArgLysSerSerSerGlyProValTrpSer 318
Qy      468 GCAGAGAGCGGTGAACCTGATATAAGACACCACCATCTCTGGGGTTCGACCCATAT 409
Db      319 GlyAlaSerSerAla-----323
Qy      408 AGATGACCTTGGAGTCTCCATTGTGGAGGTGAGAGGCGAGAAAGCGGTAGGCGAGTGGC 349
Db      323 -----323
Qy      348 TTTTATGAGCGGGGAGAGGTGAGTTCCTTGTATGATGTCAGGCGCGGTGTCGGTACT 289
Db      324 -----CysThrSerProAlaValGlyLys 331
Qy      288 CCAGGA-----CCGGGAGCTGCTGCTCGATGTTTCATCAAGCCGA 250
Db      332 AlalysSerSerGluPheAlaSerIleProAlaAsnSerSerArgProLeuSerAsnIle 351
Qy      249 TCTCATCGGGTCTAG-----CGCCTGGCTCACCAGTAGACCACT 208
Db      352 SerLysSerGlyArgMetGlyProAlaMetLeuArgProAlaLeuProAlaGlyProVal 371
Qy      207 CTTGCAGCAAGCTGTGTCGGACTTGGGGTAGTGTAGCAGTCCACACGTCGCTGGCGCGCA 148
Db      372 GlyAlaSerSerTrpGlnAlaLys-----ArgValAspValSerGluLeuAla 387
Qy      147 CCGGAAAGTTGGCATCTCTCCATCTTCCCGGAGAGGCGGAGCGGCGGCGCAGCCAT 88
Db      388 AlaGluGlnLeuThrAlaProProSerAlaSerProThrGlnProGlnThrProGluGly 407
Qy      87 GGA-----ACTCGAAGTACTTGTCTCGAACTCCCGGGGTGCTGGGGTCTCGG 37
Db      408 GlyGlyGlnTrpLeuAsnSerSerCysAla-----TrpSerGluSer 421
Qy      36 CCTGCTCTCCGCGCATGCCCGCGCGCTGCGCGTCCG 1
Db      422 SerGlnLeuAsnLysThrArgSerIleArgArgArg 433

RESULT 70
US-10-199-229-9
; Sequence 9, Application US/10199229
; Publication No. US20060099701A1
; GENERAL INFORMATION:

```







```
Db 1039 GlyProSerSerArgSerProAlaThrThrLeuProLysArgProThrSerIleLysThr 1058
Qy 543 AGGAGCGTAGCCAGCTTATCATTAACCTCCGCGCAGAAATCTTGAAGTGCTC 484
Db 1059 GluGlyLys---ProAlaAspValLysArgMetThrAlaLysSerAlaSerAlaAspLeu 1077
Qy 483 GGTAGCTCGTCCGCGAGAGCGGTGGAACTGATAATAGACACACACACACCTTGG 424
Db 1078 SerArgSerLysThrThrSerAlaSerSerValLysArgAenThrPro----- 1094
Qy 423 GGTTCGAGCCATATAGATGACCTTGGAGTCTCCATTGTGGAGTTCAGAGGCGAGAAGC 364
Db 1094 ----- 1094
Qy 363 GGTAGGCGAGTGCTCTTGATAGGCGGCGAGAGTCAGTTCCTTGATG-----ATGT 310
Db 1095 -----ThrGly-AlaAlaProProAlaGlyMet-ThrSerThrArgValLysProMets 1112
Qy 309 CCAGCGCGGCTGTGGGTACTCCAGGACCGGAGCTGCTCGATGTTTCATCAAGCGA 250
Db 1112 eAlaPro-----SerArgSerSerGlyAlaLeuSerValAspLysProT 1128
Qy 249 TCTCA-----TCGGGGTCAGCGCCCTGGCTCACCAAGTAGACACACCTCTCGACGA 199
Db 1128 hrSerThrLysProSerSerAlaProArgValSerArgLeuAlaThrThrValSer- 1147
Qy 198 AGCTGGTCCCGACTTGGGTAGGTAGGACGATCCACAGCTCCTGGCGGACCGGAGT 139
Db 1148 -----AlaProAspLeu-----LysSerValArgSerLysValGlySerThrGluAenI 1164
Qy 138 TGGCGATCTCCTCCATCTTCCCGCGCAGAGGCGGCGGCGGCGGCGGCGGCGGCGG 79
Db 1164 le-----LysHisGlnProGlyGlyArgAlaLysValGluLysL 1178
Qy 78 AGTACTTGCTCTCGAACTCCCGCGG-----GTGCTGGGGGTCTGG 37
Db 1178 ySerThrGluAlaAlaThrThrAlaGlyLysProGluProAsnAlaValThrLysAlaAlaG 1198
Qy 36 CTGCTCTC-----TCGCGCATGCGCGCGCGCTG 8
Db 1198 lySerIleAlaSerAlaGlnLysProProAla 1208
```

## RESULT 74

```
US-10-196-749-451
; Sequence 451, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Deenoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
```

```
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 451
; LENGTH: 1743
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-451

Alignment Scores:
Pred No.: 734 Length: 1743
Score: 85.00 Matches: 101
Percent Similarity: 31.1% Conservative: 26
Best Local Similarity: 24.7% Mismatches: 173
Query Match: 1.9% Indels: 109
DB: Gaps: 20

US-10-768-158-1 (1-2419) x US-10-196-749-451 (1-1743)
Qy 7 GCGACGGCGGCGGATCGCGAGAGAGCGGCGGAGAGACCCCGAGAGACCCCGGGGAGTTG 66
Db 933 AlaThrAlaAlaAla-----GlyThrGly 940
Qy 67 AGAGCAAGT---ACTTCGAGTTCATGGCTGGCGGCTGCCGCCCTTCTGCGCGGGAAGA 123
Db 941 GlyAlaGlyCysCysAlaCysThrCysAlaCysThrCysCysAlaGlyAlaAlaAla--- 959
Qy 124 TGGAGGAGATCGCCAACTTCCCGTGGCGGCCAGCAGCTGTGATCGTCACCTACCCCA 183
Db 960 -----AlaGlyGlyThrGlyGlyAlaThrAlaGlyAlaGlyThrGly 974
Qy 184 AGTCCGCGACAGCTTCTGTCGAGGAGGTGTTCTACTTGGTGAGCGGCGCTGACCCCG 243
Db 975 ThrThrCysAlaThrCysCys-----CysCysAlaGlyAlaThrThrThr 989
Qy 244 ATGAGATCGGCTTGAATGAACATCGACGAGCAGCTCCCGGCTCCTGGAGTACCACAGCGG 303
Db 990 ThrCysCysAlaThrThrThrCysThrGlyCysCysCysThrCysCysAlaAla 1009
Qy 304 GCCTGGACATCATCAAGGAACCTGACCTCTCCCGCCTCATCAAGAGCACCTGCCCTACC 363
Db 1010 ThrCysThrGlyGly-----AlaAlaAlaCysCysAlaThrCysCysThr 1024
Qy 364 GCTTTCGCTCTGACCTCCACAATCGAGACTCCAAGTCTATATATGCTCGCAACC 423
Db 1025 CysCysCysGlyAlaAlaGlyAla-----ThrGlyGlyGlyCysAlaThr 1039
Qy 424 CCAAGGATCTGGTGGTGTCTTATTATCAGTTCACCGCTCTCTGC-----GGACCA 474
Db 1040 -----CysCysAlaAlaAlaThrGlyCysCysThrThrThrGlyAla 1054
Qy 475 TGAGCTACCGAGGCA-----CCTTTCAAGAAATCTGCCCGA 510
Db 1055 CysAlaAlaAlaAlaAlaThrGlyCysThrGlyAlaThrThrThrThrCysThrGly 1074
Qy 511 GGTATTGAATGATGAAGCTGGGCTACGGCT-----CCTGTTGAGACGCTGCGAGG 561
Db 1075 GlyAlaAlaThrThrGlyCysAlaAlaAlaGlyAlaGlyAlaCysThrCysCys 1094
Qy 562 AGTTCTGGGAGCACCGCATGGACTCGAAGCTGCTTTTCTTCTCA----- 603
Db 1095 CysThrGlyCysAlaGlyGlyThrThrThrCysThrAlaAlaAlaGlyCysAlaAlaCys 1114
Qy 604 -----AGTATGAAGACATGTCATCGGACCTGGTGACGATGGTGAGCAGCTGG 651
```











## Alignment Scores:

Pred. No.: 6.55 Length: 623  
 Score: 83.50 Matches: 45  
 Percent Similarity: 32.7% Conservative: 7  
 Best Local Similarity: 28.3% Mismatches: 69  
 Query Match: 1.9% Indels: 38  
 DB: 6 Gaps: 6

US-10-768-158-1 (1-2419) x US-10-505-928-81 (1-623)

QY 1 GCGACGGCAGCGCGCGCGCATGGCGGAGAGAGCGGAGAGCCCGCAGACCCCGCGGGG 60  
 DB AlaAlaAlaLeuGlyProAlaGlyGluArgAlaGlySerArgProGlnSerProSerGly 317  
 QY 61 AGTTCGAGAGCAAGTACTTCGAGTTCCATGGCGTGGCGTCCCGCCCTTCTCCCGCGGA 120  
 DB 318 Asp-----AlaAspAlaArgGlyAspAlaAlaValTyrCysPheHisAlaAlaGly 335  
 QY 121 AGATCGAGGAGA-----TCGCCAATTCCTCCGGTGGCGCCAGG 159  
 DB 336 GluTrpArgGluLeuThrArgLeuProGluGlyAlaProAlaArgGlyCysGlyLeuCys 355  
 QY 160 ACGTGTGATCGTCACTACCCAGTCCGCGCAGCAGCTTGTCTGAGGAGTGGTCTACT 219  
 DB 356 Val-LeuTyrAnfYrLeu-----PheValAlaGlyGlyValAlaPr 369  
 QY 220 TGGTGAGCGGCGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGACGAGCAGCTCC 279  
 DB 369 oAlaGlyPro-----AspGlyArgAlaArgPr 378  
 QY 280 CGGTCTGGAGTACCACAGCGCGCGCTGGACATCATCAAGAACTGACCTCTCCCGCC 339  
 DB 378 oSerAspGlnValPheCysTyrAsnProAlaThrAspSerTrpSerAlaValArgProLe 398  
 QY 340 TCATCAAGACCACTGCGCTACCGCTTCTGCGCTCTGACCTCCACATCGAGACTCCA 399  
 DB 398 uArgGlnAlaArgSerGlnLeuArgLeuLeuAla-----LeuAs 411  
 QY 400 AGGTCACTATATGCTGCTGCCAACCCAGGATCTGTGTGTCTTATTATCAGTT 454  
 DB 411 pGlyHisLeuTyrAlaVal-----GlyGlyGluCysLeuLeuSerVal 425

## RESULT 82

US-10-523-014-5  
 ; Sequence 5, Application US/10523014  
 ; Publication No. US20060094101A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lin, Lin-Ling  
 ; APPLICANT: Yannoni, Yvonne  
 ; TITLE OF INVENTION: MK2 Interacting Proteins  
 ; FILE REFERENCE: 08702-0097-00304  
 ; CURRENT APPLICATION NUMBER: US/10/523,014  
 ; CURRENT FILING DATE: 2005-02-01  
 ; PRIOR APPLICATION NUMBER: USN 60/400,044  
 ; PRIOR FILING DATE: 2002-08-02  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 5  
 ; LENGTH: 433  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-523-014-5

## Alignment Scores:

Pred. No.: 6.25 Length: 433  
 Score: 83.00 Matches: 58  
 Percent Similarity: 28.2% Conservative: 14  
 Best Local Similarity: 22.7% Mismatches: 61  
 Query Match: 1.9% Indels: 122  
 DB: 6 Gaps: 11

US-10-768-158-1 (1-2419) x US-10-523-014-5 (1-433)

QY 714 GGTGGCAGTGTCTCGTCAAGGCTTCCAGTGGGCTTGTCTCAGAGACACCCCGAGGAATC 655  
 DB 5 GlyGlyCysSerProArgAlaProAlaAlaProGlnProArgProProProAlaLeu 24  
 QY 654 TGGCAGCTGTCTCCACCATCTGTCCAGGTCCTCCAGTGCATGCTCTTCATCTTGCAGAAAA 595  
 DB 25 ProProArg-----ProArgAlaProValPro----- 33  
 QY 594 GCACGTTTCGAGTCCATCGGTCTCC-----AGAACTCTCTGCACGTGTCAAACACGAGGAGC 538  
 DB 34 -----AlaSerArgProGlyArgProLeuLeuThrProAlaArgProCysGlyArgMet 51  
 QY 537 CGTAGCCAGCTTATCATTAACCTCCCGCAGAAATCTTTGAAAGGTGCTCGGTAGC 478  
 DB 52 Arg----- 52  
 QY 477 TCATGGTCCGAGAGCGGTGGAACGTATAAAGACACCCACCATCTCTTGGGTTCG 418  
 DB 53 -----ArgGlySerProGlyPro----- 58  
 QY 417 GAGCATATAGATGACCTTGGAGTCTCCATTTGTGAGGTTCAGAGGCGAGAAAGCGGTAGG 358  
 DB 59 -----ArgLeuGlyGlySerArgGlyGluArgArg 69  
 QY 357 GCAGGTGGCTTTCATGAGGCGGAGAGGTCACTTCTTGTATGATGATGATGATGATGATGAT 298  
 DB 70 ProAlaGlyArg-----AspProAla 76  
 QY 297 GTGGGTACTCAGACCGGAGCTCGTCTCGATGTTTCATCAAGCCGATCTCATCGGGT 238  
 DB 77 ArgValGlyProGlyGlnGlyLeuArg 86  
 QY 237 CAGCGCCCTGCTCACCAGTAGACACCTCTCTGCAGCAAGCTGTGCGGACTTGGGGT 178  
 DB 87 -----ProAlaArgProGlyProAlaAla---TrpThrGluThr-GlyG 100  
 QY 177 AGGTGAGATCCAC-----ACGTGCTGGGCGGACCGCGGAAGTTG----- 137  
 DB 100 nGly-11eValHisAlaLeuThrAspLeuSerIleProGlyMetThrSerGlyAsnGlyA 120  
 QY 136 --GGATCTCTCTCATC----- 122  
 DB 120 snSerAlaSerSerIleAlaGlyThrAlaProGlnAsnGlyGluAsnLysProProGlnA 140  
 QY 121 -----TTCCCGCGGAGAGGCG 103  
 DB 140 lAlleValLysProGlnIleLeuThrHisValIleGluGlyPheValIleGlnGlyA 160  
 QY 102 GCAGCCGAGCCCATGAACTCGAAGTACTTGTCTCTCGAAC 62  
 DB 160 laAspValSerArgTrpAspAlaArgLeuLeuValGlyAsn 173

## RESULT 83

US-11-311-754-31  
 ; Sequence 31, Application US/11311754  
 ; Publication No. US20060099635A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Vincent Ling  
 ; APPLICANT: Kyriaki Dunussi-Joannopoulos  
 ; TITLE OF INVENTION: NOVEL GLU5 MOLECULES AND USES THEREFOR  
 ; FILE REFERENCE: GNN-007  
 ; CURRENT APPLICATION NUMBER: US/11/311,754  
 ; CURRENT FILING DATE: 2005-12-19  
 ; PRIOR APPLICATION NUMBER: US/09/667,135  
 ; PRIOR FILING DATE: 2000-09-21  
 ; NUMBER OF SEQ ID NOS: 38  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 31  
 ; LENGTH: 558  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:





QY 352 ACCTGCCCTA-----CGCTTTCTGCGCTCTGACCTCCACAATGG---AGACTCCAAGG 402  
Db 208 gProArgLeuTyrglyProGlyGlyAspLeuLeuProProAlaTrpLeuArgAlaLeuG 228  
QY 403 TCATCTTA-----TATGGCTCGCAACCCCAAGGATCTGGTGGTGTCTTATTATC 450  
Db 228 yAlaLeuLeuLeuAlaLeuSerAlaLeuPheSerGlyLeuArgLeuSerLeuSe 248  
QY 451 AGTTCCACCGCTC-----TCTGGGACCATGAGCT 480  
Db 248 rLeuAspProValGluLeuArgValLeuArgAsnSerGlySerAlaAlaGluGlnGlu 268  
QY 481 ACCGAGCACCTTTCAAGAACTCTCCCGAGGTTTATGAATGATAAGCTGGCTACGGCT 540  
Db 268 nAlaArgValGlnAlaValArgGlyArg-----GlyThrHisLe 282  
QY 541 CCTGGTTTGACAGCTGCGAGGTTCTGGGAGCACCGCATGGAATCGAAGCT---GCTTT 597  
Db 282 uLeuCysThrLeuLeuLeuGlyGlnAlaGlyAlaAsnAlaAlaLeuAlaGlyTrpLeuTy 302  
QY 598 TTCTCAAGTATGAACATGATCGGGACCTGGGACCTGGTGACGATGGTGAGAGCTGGC 652  
Db 302 rAlaSerLeuProGlyValGlyAspProGlyGluAspSerGlyGluAlaGly 320

## RESULT 87

US-10-196-749-518  
; Sequence 518, Application US/10196749  
; Publication No. US20060094864A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C340  
; CURRENT APPLICATION NUMBER: US/10/196,749  
; CURRENT FILING DATE: 2002-07-16  
; PRIOR APPLICATION NUMBER: 10/052586  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059266  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063120  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063121  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063486  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/063540  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063541  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063544  
; PRIOR FILING DATE: 1997-10-28  
; Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 518  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-196-749-518

Alignment Scores:  
Pred. No.: 9.11 Length: 1137  
Score: 83.00 Matches: 170  
Percent Similarity: 28.7% Conservative: 84  
Best Local Similarity: 19.2% Mismatches: 292  
Query Match: 1.9% Indels: 340  
DB: 6 Gaps: 38  
US-10-768-158-1 (1-2419) x US-10-196-749-518 (1-1137)  
QY 2262 TTTTCACAAAGAGCAGCTCCCTCACAGAAAACACAGCAAGTGTGCGGAGCCCAACATG 2203  
Db 280 TTPAsnArgSerGluGlyTrpSerArgAsnHisArgAlaGlyGlyProGlu----- 296  
QY 2202 CGATGCCACGTGCGACCTGCTGCGCCACGCTTCTCCCATTTACAAACAAAACAACAGCT 2143  
Db 297 -----ArgProSerTrpProArgArgAspHisSerArgAsnSerAlaGlnGly 312  
QY 2142 TCTCTCTACACCCGTTTCAAAATACAGCACCAAGACGAGACCCCTTCCACAGCAGCGAC 2083  
Db 313 SerPheLeuValSerVal-----AlaGluProGlyThrProAlaIleLeuSerSer 329  
QY 2082 CCTCCTCACCCACTTTGGGTGCTTTCCAGCCACATAAATGACAGACAAACCCCGGTCCC 2023  
Db 330 ProGluPheGlnAlaSerGly-----ThrSerAsnCysSerLeu 342  
QY 2022 CAGGGGTGGTGGTACTGCGCCCGACTCTCCATCGCTTACGCGAATGTGGGTAGA 1963  
Db 343 ValPheTyrglnTyrlLeuSerGlySerGluAlaGlyCysLeuGln---LeuPheLeuGln 361  
QY 1962 AAATTAGGCAGTGAAGCAATAAAGACCTCACCTTACTCTCCCTCAGATACACGA----- 1909  
Db 362 ThrLeuGlyProGlyAlaProArgAlaProValLeuLeuArgArgArgGlyGluLeu 381  
QY 1908 GGGATGAGGTGG-----ACAGACACGAGCGCGCGCTGCCAACCTGATCCGAG 1861  
Db 382 GlyThrAlaTrpValArgAspArgValAspIleGlnSerAlaTyPro----- 397  
QY 1860 TGAACAGGGCTACACTCGCAAAATGTCTCTCCACGGCTGAGCGCTCAGGTGCACAGA 1801  
Db 397 ----- 397  
QY 1800 ACAAACTGAATGATAAGACCCCGGGATTTTAGAAATTTACTATTAAAGCATCTGTTGCC 1741  
Db 397 ----- 397  
QY 1740 AGCATTAGTAGGCTATCATCTTCAGGAAACAAAGTAAGGCTGTGGAAATTCAGTCAAATTC 1681  
Db 398 -----PheGlnIleLeuLeu 402  
QY 1680 CGAGCGCAACTCGGTCTCTTTGGTGGC-----CGACTCTCCACCC 1642  
Db 403 AlaGlyGlnThrGlyPro--GlyGlyValValGlyLeuAspAspLeuIleLeuSerAspH 422  
QY 1641 ACTGCATGCAAAATGCTCTGACTGCATTAGACTCCACCTGCCAGCTCACCAGCGCAGAC 1582  
Db 422 iCysArgPro-----ValSerGluValSerThrLeuGlnProLeuProGly---- 438  
QY 1581 TTCCCTCAGCACATGCTCCGCGGCTCTCTGCAATGGGCTTTGGAGCGCCCGCT--- 1525  
Db 439 -----ProArgAlaPro-AlaProGlnProLeuProProSerSerArg 452  
QY 1524 -----GGAAGTCTCTGGCAGAGGACGTCATCTCGAAGACCTAGTG-----AAC 1480  
Db 453 LeuGlnAspSerCysTysGlnGlyHisLeuAlaCysGlyAspLeuCysValProGlu 472  
QY 1479 AATTTCTGCTCATTTTCAGGAA----- 1459  
Db 473 GlnLeuCysAspPheGluGluGlnCysAlaGlyGluAspGluGlnAlaCysGlyThr 492  
QY 1458 -----GGTTCTGAAGAGTGAACCTCGTCAGAAATAAC 1426  
Db 493 ThrAspPheGluSerProGluAlaGlyGlyTrpGluAspAlaSerValGlyArgLeuGln 512



```
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 97
; LENGTH: 1894
; TYPE: PRP
; ORGANISM: Homo Sapien
US-10-196-749-97

Alignment Scores:
Pred. No.:      11.1      Length:      1894
Score:          83.00     Matches:      83
Percent Similarity: 32.3%  Conservative: 29
Best Local Similarity: 23.9% Mismatches:    150
Query Match:     1.9%     Indels:       85
DB:              6       Gaps:         16

US-10-768-158-1 (1-2419) x US-10-196-749-97 (1-1894)

QY 61 AGTTGAGAGCAAGTACTTCAGTTCCATGCGTGC---GGCTGCGGCCCTTCTGCGCG 117
D 235 ThrThrAlaThrThrGlyAlaAlaAlaAlaAlaCysThrAlaCysAlaGlyCysThrGly 254
QY 118 GGAAGATGGAGAGATGCCAACTTCCCGGTGCGGCCGAGCAGCGTGTGGATCGTCACCT 177
D 255 GlyAlaGlyAlaThrAlaThrThrGlyAlaCysAlaThrAlaGlyAlaGlyThrThrGly 274
QY 178 ACCCAAGTCGGCACCAGCTTGTGCGAGGAGTGTCTACTTGTGTGAGCCAGCGCGGTG 237
D 275 Thr-----GlyGlyThrCysCysAlaAlaAlaGly-----AlaAlaGlyCys 288
QY 238 ACCCGATGACATCGCTTGTGAACATCGACGAGCAGCTCCCGTCTCGGTGAGTACCCAC 297
D 289 ThrCysCysThrAlaAlaAlaGlyCysThrThrGlyCysAlaGlyAlaAlaAlaThr--- 307
QY 298 AGCCGGGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGC 357
D 308 -----ThrThrThrAlaThrCys 313
QY 358 CCTACCGCTTCTGCCCTCTGACCTCCACATGGAGACTCCAAGTCACTATATGGCTC 417
D 314 ---CysAlaAlaCysThrThrThrGlyThrThrGlyGlyAlaAlaGlyCysThrThr 332
QY 418 GCAACCCCAAGATCGTGTGTCTTATTATTCAGTCCACCGCTCTCTGCGGACCATGA 477
D 333 AlaThrThrAlaThrGlyAlaCys-----AlaAlaThrAlaCysCysAlaThrThr 349
QY 478 GCTACCGAGGACCTTCAAGAATTTCTCCGAGGTTTTATGAATGATAGCTGGGCTACG 537
D 350 ThrThr-----ThrCysAlaThr 355
QY 538 GCTCTGTTTGAGCAGCTGAGGAGTTCTGGGAGCACCAGTACGACTCGAAGCTGCTTT 597
D 356 Ala---GlyAlaGlyThrThrGlyThrGlyCysCysThrGlyGlyThrThrCys--- 373
QY 598 TTCTCAAGTATGAACATGATCGGACCTGGTGTGACGATGGTGAGAGCGTGGC----- 652
D 374 ---AlaThrAlaGlyThrCysCysAlaAla---GlyGlyCysGlyGlyAlaGlyAlaThrCy 392
QY 653 -----CAGATTCTGGGGTGTCTGTGACAAAGGCCAGCTGGAAGCCCTGAGCGAGC 705
D 392 sCysThrAlaCysThrGlyGlyCysAlaCysAlaGly---GlyGlyAlaGlyThrGlyGl 411
```

```
QY 706 ACTGCCACACAGCTGTGGACCACTGCTGCAACGCTGAGGCCCTGCGCGTGGCGGGGAA 765
D 411 yThrGlyGlyAlaGlyAlaGlyThr-CysThrAlaThrCysThrAlaThrGly-GlyAla 430
QY 766 GAGTTGGCTGTGGAGGAGCATCTTCACCGCTCTCCATGAATGAGAACTTTGACTTGTGT 825
D 431 GlyCysGlyCysCysAlaThrThrCysAlaAla-----AlaGlyAlaThr-Gly-- 446
QY 826 ATAAACAGAAAGATGGGAAAGTGTGACCTCAGCTTGTGCTTTTATTATAATAACAGAAAC 885
D 447 -----AlaAlaTh 449
QY 886 AACAACTGCTGCTCACAATACCCAGACAGTCTACTAGCCAAAAGTCTGTATGCATTTC 945
D 449 rThrThrCysAlaThrThrCysAlaCysGlyGlyThrThr----- 462
QY 946 ATTTATTCCTTGTGGACAAACTCTGGAAGCAGCG---TGTGAACACGCGGGGAAAGGA 1002
D 463 -----GlyCys-GlyThrThrThrAlaAlaThrCysGlyGlyAlaGlyAlaGly- 479
QY 1003 AGAGCGCGGTGAGCGGAGGAGTGTGATGATTCCTCAACCGAAGCAGCTGTCTGCCTTTA 1062
D 480 --GlyAlaCysThrGlyGlyThrThrGlyCysCysAlaThrGlyGlyCysAlaAlaAla 499
QY 1063 GAACGTGCAGC 1073
D 499 hrGlyCysThr 502

RESULT 89
US-10-370-959-40
; Sequence 40, Application US/10370959
; Publication No. US20060088907A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williams, Mark W.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Hunter, John J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Bandaru, Rajasekhar
; APPLICANT: Tsai, Fong-Ying
; TITLE OF INVENTION: NOVEL 13237, 18480, 2245, 16228, 7677,
; TITLE OF INVENTION: 26320, 46619, 33166, 18836, 48867, 21617, 55562, 39228,
; TITLE OF INVENTION: 62088, 46745, 23155, 21657, 42755, 32229, 46863 AND
; TITLE OF INVENTION: 32252 MOLECULES AND USES THEREFOR
; FILE REFERENCE: MPI03-0170NMIM
; CURRENT APPLICATION NUMBER: US/10/370,959
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: US 09/910,150
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/219,028
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 10/251,507
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US 09/715,479
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US 60/218,053
; PRIOR FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: US 09/644,929
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 60/212,439
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 09/892,870
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: US 60/214,174
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: US 09/775,117
; PRIOR FILING DATE: 2001-02-01
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: FastSeq for Windows Version 4.0
```



```
; SEQ ID NO 40
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-370-959-40

Alignment Scores:
Pred. No.:      8.14      Length:      665
Score:          82.50     Matches:      60
Percent Similarity: 32.6%   Conservative: 31
Best Local Similarity: 23.8% Mismatches: 104
Query Match:    1.9%     Indels:      75
DB:             6        Gaps:       13

US-10-768-158-1 (1-2419) x US-10-370-959-40 (1-665)
QY 789 AGATGCTCTTCACAGCCAACTCTCCCGCCCGCCAGCGAGGCCT-----CAGCGT 736
Db 19 GlnCysProValCysGlnGlnMetMetProAlaAlaHisIleAsnSerHisLeuAspArg 38
QY 735 TGCAGCACTGGTCCACAGCTGGTGGCAGTGTCCGTCAGGCGCTTCCAGCTGGCGCTTGT 676
Db 39 CysLeuLeuLeuHisProAlaGlyHisAlaGlu-----ProAlaAlaGlySer 54
QY 675 CACAGGACACCCCGAGGAATCGGCCAGCTGTCCACCATCGTCACAGGTCCTCCATGCA 616
Db 55 HisArgAlaGlyGlu-----ArgAlaLysGlyProSerProGly-----68
QY 615 TGTCTTCATCTTGAGAAAAGCAGCT-----TCAGATCCATCGGTGCTCC 568
Db 69 -----AlaLysArgArgLeuSerGluSerSerAlaLeuLysGlnPro 83
QY 567 AGAACTCTGCAGCTGCTCAACACGAGCCGCTAGCCCGCTTATTCATATAAACCTCC 508
Db 84 AlaThrProThrAlaAlaGluSerSerGlu-----93
QY 507 GCGAAGATCTTGAAGGTGCTCGTAGCTCATGTCGCGAGAGCGGTGGA---ACT 451
Db 94 -----GlyGluGlyGluGlyAspAspGlyGlyGluThr 105
QY 450 GATAATAAG-----ACACACAGATCCTTGGGGTGGCGCCATATAGATGA 403
Db 106 GluSerArgGluSerTyrAspAlaProProThrProSerGly-AlaArgLeuIleProAs 125
QY 402 CTTGGAGTCTCCATGTGGAGGTGAGAGGACGAGAAACGGTAGGCGAGGTGGCTCTTGA 343
Db 125 pPheProValAlaArgSerSerSerProGlyArgLysGlySerGlyLysArgProAlaAl 145
QY 342 TGAGCGGGGAGAGT-----CAGTCTTGATGATGTCAGGCGCGGCTGTG 295
Db 145 aAlaAlaAlaGlySerAlaSerProArgSerTrpAspGluAlaGluAlaGln-----163
QY 294 GGTACTCTCAGACCGGAGCTGCTCGTAGTGTTCAT-----CAAGCGCATCTCAT 244
Db 164 -----GluGluGluGluAlaValAlaGlyAspGlyAspGlyAspAlaAspAlaAs 181
QY 243 CGGGGTACGCGCTCGGTCCACCAAGTAGACACCTCTCGCAGCAAGCTGTGCC---GG 187
Db 181 pGlyGluAspAspProGlyHisTrpAspAlaAspAlaAlaGluAlaAlaThrAlaPheGl 201
QY 186 ACTTGGGTGAGTACGATCCACAGTC---GCTGGCGCGCAGCCGGGAAGTTGGGATCT 130
Db 201 yAlaSerGlyGlyArgProHisProArgAlaLeuAlaAlaGluGluIleArgGlnMe 221
QY 129 CTTCCAT---CTTCCCGCGGAGAGGCGCGCAGCCGCGCAGCCATG-----87
Db 221 tLeuGlnGlyLysProLeuAlaAspThrMetArgProAspThrLeuGlnAspTyrPheGl 241
QY 86 -----GAACT 82
Db 241 yGlnSerLysAlaValGlyGlnAspThrLeuLeuArgSerLeuLeuGluThrAsnGluIl 261
QY 81 CGAAGTACTTCTGCTCGAACTCCCCCGGGGTGCTGGGGGTCTCGCGCTCGCT 30
```

```
Db 261 eProSerLeuIleLeuTrpGlyProProGlyCysGlyLysThrThrLeuAla 278
RESULT 90
; Sequence 393, Application US/10505928
; Publication No. US20060089532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 393
; LENGTH: 667
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-393

Alignment Scores:
Pred. No.:      8.14      Length:      667
Score:          82.50     Matches:      60
Percent Similarity: 33.7%   Conservative: 31
Best Local Similarity: 22.2% Mismatches: 104
Query Match:    1.9%     Indels:      75
DB:             6        Gaps:       13

US-10-768-158-1 (1-2419) x US-10-505-928-393 (1-667)
QY 723 CCACCA-----GCTGTGGCAGTCTCCGTCAGGGCTTCAGCTGGGCTTGTGCAC--- 673
Db 28 ProProValSerSerGlyLysAsnGlyProThrSerLeuAlaSerGly-----HisPhe 45
QY 672 -----AGGACACCCCGAGGAATCTGGCCAGCTGTCCACCATCG 634
Db 46 ThrGlySerAsnValGluAspArgSerSerGlySerTrpGlyAsnGlyGlyHisPro 65
QY 633 TCACCAAGT-----CCCGATGCATGCTTTCATCTTGAGAAAAA 595
Db 66 SerProSerArgAsnTyrGlyAspGlyThrProTyrAspHisMetThrSerArgAspLeu 85
QY 594 GCAGTTCGAGTCCATCGGTGCTCCAGAACTCTGCACGCTGTCTCAAACAGGAGCGGT 535
Db 86 GlySerHisAspAsnLeuSerProPheValAsnSerArgIleGlnSerLysThrGlu 105
QY 534 AGCCAGCTTATCATTAACCTCCGGCAGAAATCTTGAAGGTGC-----487
Db 106 ArgGlySerTyrSerSerTyrGlyArgGluSerAsnLeuGlnGlyCysHisGlnGlnSer 125
QY 486 ---CTCGGTAGTTCATGTCGCGCAGAGCGGTGGAACTGATAATAGACACCAACCATG 430
Db 126 LeuLeuGlyGlyAspMetAspMetGlyAsnProGlyThrLeuSerProThrLysPro---144
QY 429 CTTTGGGTTTCGAG-----CCATATAGATGACCTTGGAGT 394
Db 145 -----GlySerGlnTyrTyrGlnTyrSerSerAsnAsnProArgArgArgPro-----160
QY 393 CTCATTGTGGAGGTTCAGAGGGCAGAAAGCGGTAGGCGAGGTGGCTCTTGTATAGGCGGG 334
Db 161 LeuHisSerSerAlaMetGluValGlnThrLysLysVal-----173
QY 333 GAGAGGTTCAGTTCCTTCATGATGTCCAGGCGCGGCTGTGGGTACTCCAGGACCGGGAGCT 274
Db 174 ---ArgLysValPro-----ProGlyLeuProSerSerValTyrAlaProSerAla 189
QY 273 GCTCGTCGATGTTTCATCAAGCCGATCTCATCGGGGTTCATCGGGGTTCAGGAGTGTAG 214
Db 190 -----SerThrAlaAspTyrAsnArgAspSerProGlyTyrProSerSer 204
```

213	CCACCTCTCTGCAGCAAGCTGGTGGCCGGA	CTTTGGGTAGGTGACGATCCACACCTCGCTCG	154
		:::>:::>::	
205	LysProAlaThr-SerThr-Phe--Pro-		212
153	GCGCACACGGGAAGTTGGCGATCTCCTCCATCTTCCGCGGCAGAGGGCGGACGCGCA	94	
		:::>::	
213	-----SerSerPhePheMetGlnAspGlyHisHisSerSerA	225	
93	CGCCATGGAACTCGAAGTACTTGCTCTCGAACTCCCGGGGTGCTGGGGGTCTCGCGCT	34	
	>:::>::		::
225	spProTrpSerSerSerGlyMet----	AsnGlnProGlyTyrAlaGlyMetLeuGlyA	244
33	CGCTCTCCGCCATGCGCGCGCGTGC	8	
	::		
244	snSer-SerHisIleProGlnSerSer	252	
	::		

```

RESULT 91
US-10-196-749-97
; Sequence 97, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCES: P34301C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612

```

US-10-768-158-1 (1-2419) x US-10-196-749-97 (1-1894)

Qy 759 GGCCACGGGCGAGGCCCTCAGCGTTGCAGCACTGGTCCACCA---GCTGGTGCGAGTCGT 703  
Db 1019 GlyAlaalaAlaGlyAlaAlaCysCysThrGlyAlaGlyAlaGlyAlaGlyAla 1038  
Qy 702 CCGTCAGGCTTCAGCTGGGCCCTTGTCAAGCACCCCAGGAATCTGGCCAGCTGCT 643  
Db 1039 AlaAlaGlyAlaAlaThrThrGlyCysCysAlaAlaAlaAlaAlaThrThrAlaAla 1058  
Qy 642 CCACCATCGTCACCAAGCTCCCGATGCATGCTCTCATACTTGGAAAAAGACCGTTCAGT 583  
Db 1059 AlaAlaAlaAlaGlyAlaCysAlaCys-----AlaAlaGlyThr 1072  
Qy 582 CCATGCGGTGCTCCAGAACTCTCTGCAGCTCTCAAACCCAGAGCCGTAGCCCCAGCTTAT 523  
Db 1073 GlyCysGlyAlaAlaThrThrGlyThrAlaAlaAlaThr----- 1085  
Qy 522 CATTCATAAACCTCCGCGAGAATTCTTGAAGGTGCTCGGTAGCTCATCGTCCGACAG 463  
Db 1086 -----CysAlaGlyCysThrGlyGlyAlaGly 1094  
Qy 462 AGCGGTGGAATGATAATAGACACCAACCACTCTTGGGTTGCGAGCCATATAGATGA 403  
Db 1095 AlaAlaGly----- 1097  
Qy 402 CTTGGAGTCTCCATTGCGAGGTCAGAGGCGAGAAAGCGTAGGCGAGTGGCTCTTGA 343  
Db 1098 -----GlyAlaGlyAlaAlaGlyThrGlyGlyAlaGlyAlaAla--- 1110  
Qy 342 TGAGGCGGGAGAGTCACTTCTTGATGTTCAGGC-----CCGGCTGTGGGTACT 289  
Db 1111 --GlyAlaAlaThrCysAlaGlyThrCysAlaGlyCysCysGlyCysAlaGlyThr 1129  
Qy 288 CCAGACGGGAGTCTGCTGATGTTTCATCAAGCGATCTCATCGGGTCAGCGCCCT 229  
Db 1130 GlyAlaAlaGlyAlaGlyCysThrCysAlaGlyAla----- 1141  
Qy 228 GGCTCACCAAGTAGACCACTCTCTGCAGCAAGCTGGTGCCGACTGGGGTAGGTGACGA 169  
Db 1141 ----- 1141  
Qy 168 TCCACAGTCTGCTGGCGCGCACCGGAAGTTGGCGATCTCTCATCTTCCCGCGCAGA 109  
Db 1142 -----AlaAlaAlaGlyAlaAlaGlyCysAlaAlaGlyAlaCysAlaAla 1156  
Qy 108 AGGCGGCGAGCGGACCGCATGGAATCTGTAAGTACTTCTCTGAACCTCCCCCGGGTGC 49  
Db 1157 ThrThrAlaAlaAlaAlaCysGlyGlyGlyAlaAlaCysThrThrThrAlaGlyCys 1176  
Qy 48 TGGGGGTCTCGGCTCGCTCTCCGCCATCGCGCGCGCTGCGCC 6  
Db 1177 -AlaGlyCysAlaAlaAlaAlaCysAlaAlaAlaAlaAla 1190

RESULT 92  
US-11-118-524-2  
; Sequence 2, Application US/11118524  
; Publication No. US2006008847A1  
; GENERAL INFORMATION:  
; APPLICANT: GU, WEI  
; TITLE OF INVENTION: ARF-BP1 AS MEDIATOR OF P53-DEPENDENT AND INDEPENDENT TUNING  
; FILE REFERENCE: 19240-497US2  
; CURRENT APPLICATION NUMBER: US/11/118,524  
; PRIOR FILING DATE: 2005-04-29  
; PRIOR APPLICATION NUMBER: 60/610,506  
; PRIOR FILING DATE: 2004-09-15  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 2  
; LENGTH: 4373  
; TYPE: PRN  
; ORGANISM: homo sapiens

US-11-118-524-2									
Alignment Scores:									
Pred. No.:	16.8	Length:	4373						
Score:	82.50	Matches:	88						
Percent Similarity:	35.8%	Conservative:	34						
Best Local Similarity:	25.8%	Mismatches:	119						
Query Match:	1.9%	Indels:	100						
DB:	7	Gaps:	20						
US-10-768-158-1 (1-2419) x US-11-118-524-2 (1-4373)									
Qy	832	TGTTATACACCAAGTCAAACTTCATTCATGAGCGGTGAAG-----ATG	785						
Db	2721	CysThrAlaSerLysSerAsnAspSerThrGluGlnAsnLeuSerAspGlyThrProMet	2740						
Qy	784	TCCTTCCACAGCCCAACTCTTCCCGGCCCGCCAGGCGCTCAGCG-----737							
Db	2741	ProAspSerTyProThrThrProSerSerThrAspAlaIleThrSerGluSerLysGlu	2760						
Qy	736	-----TTGCAGCAGCTGGT-----CCACGAGCTGG	713						
Db	2761	ThrLeuGlyThrLeuGln-SerSerGlnGlnProThrLeuProThrProAlaLe	2780						
Qy	712	TGGCAGTGTCCGTACGGCTT-----CCAGCTGGGCTTGTACAGACACCCCGAG	659						
Db	2780	uGlyGluValProGlnGluLeuGlnSerProAlaGlyGluGlyGlySerThrGlnLe	2800						
Qy	658	AATCTGGCCAGCTGTCCA-----CCATCGTCACAGGTCCCGATCATGTCT	611						
Db	2800	uLeuMetProValGluProGluGluLeuGlyProThrArgProSerGlyGluAla----	2818						
Qy	610	TCATACTTGAGAAAAAGCAGTTCAGTCCATGCGTCTCCCGA-----AC	563						
Db	2819	-----GluThrGlnMetGluLeuSerProAlaProThrIleThrSerLeuSe	2835						
Qy	562	TCCTGCAGTGTCTAAACAGAGCGGTAGCCAGCTTATCATTAACAACTCCGGCAG	503						
Db	2835	rProGluArgAlaGluAspSerAspAlaLeuThrAlaValSerSerGlnLeuGluGlySe	2855						
Qy	502	A-----ATTCTGAAAGTGCTCGGT-----481							
Db	2855	rProMetAspThrSerSerLeuAlaSerCysThrLeuGluGluAlaValGlyAspThrSe	2875						
Qy	480	-AGCTCATGGTCCGAGAGCGGTGGAACATGATAAAGACACACACAGAT-----CC	428						
Db	2875	rAlaAlaGlySerSerGluGlnProArgAlaGlySerSerThrProGlyAspAlaProPr	2895						
Qy	427	TTGGGTTGCGAGCCATATAGATGACCTTGGAGTCTCCATTGTGGAGTCCAGAGGCAGA	368						
Db	2895	oAlaValAlaGlu-----Va	2900						
Qy	367	AAGCGGTAGGCGAGGTGCTTTGATAGGCGGGAGAGGTCAAGTTCCTTGATGATGCC	308						
Db	2900	lGlnGlyArgSerAspGlySer-----GlyGluSerAlaGlnPro-----Pr	2914						
Qy	307	AGGCCCGCTGTGGTACTCCAGGACCGGAGCTGCTCGTATGTTTCATCAAGCCGATC	248						
Db	2914	oGluAspSer-----SerProAlaSerSerGluSerSerThrArgAs	2930						
Qy	247	TCATCGGGGTGAGCCCTCGGTACCAAGTAGACCACTCTCTGAGCAA-----198							
Db	2930	pSerAlaValAlaIleSerGlyAlaAspSerArg-GlyIleLeuGluGluProLeuProS	2950						
Qy	197	-----GCTGTGCCGACTTGGGTAGTGACCATC	167						
Db	2950	erThrSerSerGluGluAspProLeuAlaGlyIleSerLeuProGluGlyValAspp	2970						
Qy	166	CACAC-----GTCGCTGGGCG-----CACCGGAAAGTT-----GGCGATCTCC	128						
Db	2970	roSerPheLeuAlaAlaLeuProAspAspIleArgArgGluValLeuGlnAsnGlnLeu-	2989						
Qy	127	TCCATCTTCCCGGCAGAGGGCGGACGCCCATGGCAACTCGAAGTACTTGCTC	68						

US-10-511-937-2466									
; Sequence 2466, Application US/10511937									
; Publication No. US20060088936A1									
; GENERAL INFORMATION:									
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.									
; APPLICANT: Wohlgemuth, Jay									
; APPLICANT: Fry, Kirk									
; APPLICANT: Woodward, Robert									
; APPLICANT: Ly, Ngoc									
; APPLICANT: Prentice, James									
; APPLICANT: Morris, Macdonald									
; APPLICANT: Rosenberg, Steven									
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION									
; FILE REFERENCE: 50661200104									
; CURRENT APPLICATION NUMBER: US/10/511,937									
; PRIOR FILING DATE: 2004-10-19									
; PRIOR APPLICATION NUMBER: PCT/US2003/012946									
; PRIOR FILING DATE: 2003-04-24									
; PRIOR APPLICATION NUMBER: US 10/131,831									
; PRIOR FILING DATE: 2002-04-24									
; PRIOR APPLICATION NUMBER: US 10/325,899									
; PRIOR FILING DATE: 2002-12-20									
; NUMBER OF SEQ ID NOS: 3117									
; SOFTWARE: PatentIn version 3.2									
; SEQ ID NO 2466									
; LENGTH: 362									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-10-511-937-2466									
Alignment Scores:									
Pred. No.:	7.06	Length:	362						
Score:	82.00	Matches:	65						
Percent Similarity:	26.4%	Conservative:	13						
Best Local Similarity:	22.0%	Mismatches:	82						
Query Match:	1.9%	Indels:	136						
DB:	6	Gaps:	14						
US-10-768-158-1 (1-2419) x US-10-511-937-2466 (1-362)									
Qy	29	GAGCGAGCGCGAGACCCC-----CAG	49						
Db	66	GluArgGlyAlaAspAlaThrTrpAspLeuAspLeuLeuThrAsnPheSerGlyPro	85						
Qy	50	CACCCCGGGGAGTTCGAGACAAAGTACTTCGAGTTCATGGCTGCGCGCCCTT	109						
Db	86	GluProGlyGlyAlaProGln-----ThrCysAlaLeu	96						
Qy	110	CTGCGCGGGAGATGGAGGAGATCGCAACTTCCCGGTGCGGCCAGCGCTGTGAT	169						
Db	97	AlaProSerGluAlaSerGlyAlaGlnTyProPro-----109							
Qy	170	CGTCACCTACCCCAAGTCCCGCACCACTTGTGTCAGGA-----GGTGGTCTA	217						
Db	110	-----ProGluThrLeuGlyAlaTyAlaGlyProGlyLeuValAlaGlyLeu	126						
Qy	218	CTTGTGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCAACGAGCAGCT	277						
Db	127	LeuGly-----SerGluAspHisSerGlyTrp	135						
Qy	278	CCCGGTCTCTGAGTACCCACAGCGCGGCTGGACATCATCAAGAACTGACTCTCCCG	337						
Db	136	ValArgProAlaLeuArgAlaArgAlaProAspAlaPheValGlyProAlaLeuAlaPro	155						



```
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 872
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-302-678-17

Alignment Scores:
Pred. No.: 9.95 Length: 872
Score: 82.00 Matches: 44
Percent Similarity: 32.0% Conservative: 19
Best Local Similarity: 22.3% Mismatches: 60
Query Match: 1.9% Indels: 75
DB: 7 Gaps: 9

US-10-768-158-1 (1-2419) x US-11-302-678-17 (1-872)
QY 1927 ACTCCCTCAGATACACGAGGAGTGGAGACACACGACGAGCGCGGTGCCAACCTG 1868
Db 672 SerProGluAepSer---ArgGluHisValAspArgHis--GlyCysIleValIysile 689
QY 1867 ATCCGAGT-----GAAACAGCGGTACACTCGCAAAATGGTCTCCACAGCGCTGAGCGC 1815
Db 690 ValArgSerSerSerThrGlyGlnIysAsnPheSerAlaProProAlaAlaProPro 709
QY 1814 CTCAGGTGCACAGAACTGAATGATTAAGACCCCGGGGATTTTAGAAATTTACTATT 1755
Db 710 ValGlnCysProProSerThrSer----- 717
QY 1754 AGCATCTGCTGCCAGCATTAGTAGGCTATCACTTCAGGAAACAAGTAAGCTGTGGAA 1695
Db 717 ----- 717
QY 1694 TTCAGCTCAAAATCCGAGCGCAACTCGTCTTGGTGGCGGACTCTCCACCCACTG-- 1638
Db 718 -----TrpGlnProGlnSerHisProArgGln 726
QY 1637 ---CATGCAAAATGCTCTCT-----GACTGCATTAGACTCCACCTCCAGC 1596
Db 727 GlyHisGlyThrSerProValGlyAspHisGlySerLeuValArgIleProProPro 746
QY 1595 CTCACGAGGACTTCCCTCAGACATGCTCCGCGGGTCTCTCGCATG-- 1545
Db 747 ---AlaHisGluArgSerLeuSerAlaTyGlyGlyAsnArgAlaSerMetGluPhe 765
QY 1544 -----GGCCTTTGGAGCGCGCGCTGGAAGTCTCTGGCAGAAGG 1506
Db 766 LeuArgGlnGluAspThrProGly--CysArgProProGluGlyAsnLeuArgAspSerA 785
QY 1505 AC-----GTCCATTCTGAAGACTAGTGAACAATTTCTGCTCAT 1467
Db 785 spThrSerIleSerIleProSerValAspHisGluGluLeuGluArgSerPheSerGlyP 805
QY 1466 TTCAGAAGGTTCTGAAGAAGTGAAGTCCGTCAGATAAACAAGTGT 1420
Db 805 heSerIleSerGlnSerIysGluAsnLeuAspAlaLeuAsnSerCys 820

RESULT 96
US-10-505-928-54
; Sequence 54, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; PRIOR FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 54

; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Human Sox-18
US-10-505-928-54

Alignment Scores:
Pred. No.: 7.95 Length: 384
Score: 81.50 Matches: 51
Percent Similarity: 40.1% Conservative: 8
Best Local Similarity: 34.7% Mismatches: 43
Query Match: 1.9% Indels: 46
DB: 6 Gaps: 11

US-10-768-158-1 (1-2419) x US-10-505-928-54 (1-384)
QY 31 GCGAGGCGGAGACCC-----CCACACCCCGGGGAGTTCGAGAGCAAGT 75
Db 178 AlaProGlnProProProGluProPheProAlaAlaSerGlySerAlaArgAlaPhe 197
QY 76 ACTTC-----GAGTTCCATGCGGTGCGGCTCCGCGCCCTTCTGCCGC 116
Db 198 ArgGluLeuProProLeuGlyAlaGluPheAspGlyLeuGlyLeuProThrProGluArg 217
QY 117 GGAAGATGGAG-----GAGATCGCCCACTTC-----CCGGTGGCGGCC 155
Db 218 SerProLeuAspGlyLeuGluProGlyGluAlaAlaPhePheProProAlaAlaPro 237
QY 156 AGCCAGCTGTGGTGGTCCACCTACCCCAAGTCCG---GCACCAGCTTGTCTGCAGG 208
Db 238 GluAsp-CysAlaLeuArgProPheArgAlaProTyArgAlaProThr-GluLeuSerArgA 257
QY 209 -----GGTGGTCTACTTGGTGAGCGGCGCTGACCCCGATGAGATCGGCTTGATGAAC 263
Db 257 spProGlyGlyCysTyGlyAlaPro-----LeuAlaGluA 269
QY 264 ATCAGAGCAGCTCCCGGTCTCGAGTACCCACAGCGCGG----- 304
Db 269 laLeuArgThrAlaPro---ProAlaAlaProLeuAlaGlyLeuTyTyGlyThrLeuG 288
QY 305 -----CTGGACATCATCAGGAAGTCACTCTCCCGCTCATCAAGAGCACCT----- 355
Db 288 lyThrProGlyProTyProGly---ProLeuSerProProGluAlaProProLeuG 307
QY 356 -----GCCCTACCGCTT 367
Db 307 luSerAlaGluProLeu 312

RESULT 97
US-10-511-937-2959
; Sequence 2959, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
```

```
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2959
; LENGTH: 620
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2959

Alignment Scores:
Pred. No.:          9.59      Length:      620
Score:              81.50     Matches:    27
Percent Similarity: 44.2%    Conservative: 15
Best Local Similarity: 28.4% Mismatches:   42
Query Match:        1.9%     Indels:     11
DB:                  6       Gaps:        5

US-10-768-158-1 (1-2419) x US-10-511-937-2959 (1-620)
QY 724 ACCAGTGTGCAACGCTGAGCGCTGCGCGGGCGGGAAGAGTTG---GGCTGTGGA 780
Db 514 SerSerThrGlyThrLysPheProValLysTrpAlaSerProGluValPheSerPheSer 533
QY 781 AGGACATCTTCACCGCTCTCCATGAATGAGAAAGTTTGACTTGGTGTATAAAC---AGAAGA 837
Db 534 ArgTyrSerSer-----LysSerAspValTyrSerPheGlyValLeuMet 548
QY 838 TGGGAAAGTGTGACCTCAGCTTGTGACTTTTATTATATACAGAAACAAACCTGCAT 897
Db 549 TrpGlu---ValPheSerGluGlyLysIleProTyrGluAsnArgSerAsnSerGluVal 567
QY 898 GCTCACAATACCCAGACA-----GTCTACTAGCCAAAGTCTGTATGCATTTCATT 948
Db 568 ValGluAspIleSerThrGlyPheArgLeuTyrLysProArgLeuAlaSerThrHisVal 587
QY 949 TATTCCTTGTGGCAAACTCTGGAAGCAGCGGTGTGAAACAGCGG 993
Db 598 TyrGlnIleMetAsnHisCysTrpLysGluArgProGluAspArg 602

RESULT 98
US-10-196-749-466
; Sequence 466, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
```

```
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 466
; LENGTH: 813
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-466
```

```
Alignment Scores:
Pred. No.:          10.7      Length:      813
Score:              81.50     Matches:    67
Percent Similarity: 35.2%    Conservative: 32
Best Local Similarity: 23.8% Mismatches:   77
Query Match:        1.9%     Indels:     106
DB:                  6       Gaps:        18
```

```
US-10-768-158-1 (1-2419) x US-10-196-749-466 (1-813)
QY 757 CCCACGGGCGAGGGCTTCAGCGTTG-----CAGCACTGGTCCACCAGCTGGTGGCAGTGC 704
Db 99 ProAspGlyGlnProValValLeuAlaProAsnHisThrAspHisCysHisTyrGlnGly 118
QY 703 TCCGTCTCAGGGCT-----TCCAGCTGGCGCTTGTACAGGACACACCCCGAGAAATCTGCC 650
Db 119 ArgValArgGlyPheProAspSerTrpValVal-----LeuCys 131
QY 649 AGCTGC-----TCCACCATCTCACCAGGTCCCGATGCATGTTCTTATCATCTGAGA 599
Db 132 ThrCysSerGlyMetSerGlyLeuIleThrLeuSerArgAsnAlaSerTyrTyrLeuArg 151
QY 598 AAA-----AGCACGTTTCGATGTCATGCGGTGCTCCAG 566
Db 152 ProTrpProProArgGlySerLysAspPheSerThrHisGluIlePheArgMetGluGln 171
QY 565 AAC-----TCCTGCACGTGCTCAACACGAGCGGTAGCCAGCTTATCATCTATCATATA 515
Db 172 LeuLeuThrTrpLysGlyThrCysGlyHisArgAspProGlyAsnLysAlaGlyMetThr 191
QY 514 AACCTCCGGCAGAAATCTTG-----AAAGTGCCTCGGTAGCTCATGTCGCCAGAGCG 459
Db 192 SerLeuProGlyGlyProGlnSerArgGlyArgGluAlaArgArgThrArgLysTyr 211
QY 458 GTGGAACCTGATAATA-----AGACAC----- 438
Db 212 LeuGluLeuTyrIleValAlaAspHisThrLeuPheLeuThrArgHisArgAsnLeuAsn 231
QY 437 -----CACCAGATCCTTGGGTTGCGAGCCATATAGATGACCTGGAGTCTCCATT 387
Db 232 HisThrLysGlnArgLeuLeuValAlaAsnTyrValAspGlnLeu---LeuArgThr 250
QY 386 GTGAGGTCTCAGGGCAGAAAGCGGTAGGCGAGTGGCTCTT----- 345
Db 251 LeuAspIle-----GlnValAlaLeuThrGlyLeuGluValTrp 263
QY 344 GATCAGCGGGGAGAGGTCTGCTTCTTGTATGATGTCACAGGCGCGGTGGTGA---CTC 288
Db 264 ThrGluArgAspArgSerArgValThrGlnAspAlaAsnAlaThrLeuTrpAlaPheLeu 283
QY 287 CAGGACCGG-----GAGCTGCTC----- 270
Db 284 GlnTrpArgArgGlyLeuTrpAlaGlnArgProHisAspSerAlaGlnLeuLeuThrGly 303
QY 270 ----- 270
Db 304 ArgAlaPheGlnGlyAlaThrValGlyLeuAlaProValGluGlyMetCysArgAlaGlu 323
QY 269 -----GTGATGTTTCATCAAGCCGATCTC---ATCGGGGTCTCAGCGCC---CTG 228
```

```
Db 324 SerSerGlyValSerThrAspHisSerGluLeuProIleGlyAlaAlaAlaThrMet 343
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 227 GCTCACCAGTAGACCACCTCTCGCAGCAAGCTGTCGCCGACTTGGGGTAGGTAGCGAT 168
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 344 AlaHisGluIleGlyHis-----SerLeuGlyLeuSerHisAsp 356
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 167 CCA 165
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 357 Pro 357

RESULT 99
US-10-504-120-21
; Sequence 21, Application US/10504120
; Publication No. US2006008829A1
; GENERAL INFORMATION:
; APPLICANT: Exelixis, Inc.
; TITLE OF INVENTION: MINRs AS MODIFIERS OF INSULIN RECEPTOR SIGNALING AND METHODS OF
; FILE OF INVENTION: USE
; FILE REFERENCE: EX03-003C-PC
; CURRENT APPLICATION NUMBER: US/10/504,120
; CURRENT FILING DATE: 2004-08-06
; PRIORITY APPLICATION NUMBER: 60/354,824
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/358,217
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 60/358,189
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 60/358,126
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 60/358,995
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/358,756
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/358,765
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/359,531
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/360,222
; PRIOR FILING DATE: 2002-02-26
; PRIOR APPLICATION NUMBER: 60/360,224
; PRIOR FILING DATE: 2002-02-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 1763
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-504-120-21

Alignment Scores:
Pred. No.: 14.4 Length: 1763
Score: 81.50 Matches: 165
Percent Similarity: 32.7% Conservative: 94
Best Local Similarity: 20.8% Mismatches: 286
Query Match: 1.9% Indels: 249
DB: 6 Gaps: 33

US-10-768-158-1 (1-2419) x US-10-504-120-21 (1-1763)
Qy 1880 CCGTGCCCAACCTGATCCGAGTGAACAGGGCTACACACTCGCAA---AATGGTCTCTCCACG 1824
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 727 ProCysSerValAspGlnLeuCysSerAlaLeuCysSerMetLeuSerGlyProLysThr 746
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 1823 CCGTGGCCCTCAGGTGTCACAGAACAACTGAATGATAAAGACCCGGGGATTTAGAAAT 1764
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 747 LeuGluArgLeuArgGlyAla-----ProGluGlyPheSerArg 759
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 1763 TTACTATTAAAGCATCTGC-----TTGCCAGCATTTAGGTAGGCTATCACTTCAGG 1716
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 760 ThrAspLeuHisLeuAlaValProValLeuThrAlaAlaIleSerTyrHis----- 777
||| : : : : : ||| ||| ||| : : : ||| : : :
```

```
Qy 1715 AAACAAAGTAAGGTGGGAATTCAGCTCAAAATTCGAGCGCAACTCGTCTCTTTGGTGG 1656
Db 778 -----AsnTyrLeuAspLysThrLysGlnArgGluMetValTyrCys 791
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 1655 CCGACTCTCCACCCACATGCAAAATGCTCTGACTGC----- 1617
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 792 LeuGluGlnGlyLeuIleHisArgCysAlaArgGlnCysValValAlaLeuSerIleCys 811
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 1616 -----ATTAGACTCCCA 1605
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 812 SerValGluMetProAspIleIleLysAlaLeuProValLeuValValLysLeuThr 831
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 1604 CCGCCGAGCCTCACCAGGAGACTTCC-----CTCAGCAGATGCTCC 1563
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 832 HisIleSerAlaThrAlaSerMetAlaValProLeuLeuGluPheLeuSerThrLeuAla 851
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 1562 GCGGGCTCTCT---CTGCATTGGGCTTTGGAGCGCGCTGGAAAGTCTCTG---GCAGA 1509
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 852 -ArgLeuProHisLeuTyrArgAsnPheAlaAlaGluGlnTyrAlaSerValPheAlaI 871
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 1508 AGGACGTCCATTCTGAAGACCTAGTGAA-----CAATTTCTGCTCATTTTCAGGAAG-- 1457
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 871 eSerLeuProTyrThrAsnProSerLysPheAsnGlnTyrIleValCysLeuAlaHisH 891
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 1456 -----TT 1455
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 891 sValIleAlaMetTrpPheIleArgCysArgLeuProPheArgLysAspPheValProPh 911
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 1454 CTGAAGAAGTGAACCTCGTCAGATAA----- 1428
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 911 eIleThrLysGlyLeuArgSerAsnValLeuLeuSerPheAspAspThrProGluLysAs 931
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 1427 -----ACAAGTGTCCAGATTCCCCCAACAGACCCCACTCATCTC 1387
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 931 pSerPheArgAlaArgSerThrSerLeuAsnGluArgProLysArgIleGlnThrSerLe 951
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 1386 TGTTCTCACCACCTATCTCAGAACCAAGTGAGACAGAAATGGCTGTGCTTAGATAATAA 1327
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 951 uThrSerAlaSerLeuGlySerAlaAspGluAsnSerValala-----G 966
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 1326 ACTACAAAATAAAGAGTTAAACCTGAGATCTTTTACCATTCGGGTGGTGCCTCGCTCT 1267
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 966 nAlaAspSerLeuLysAsnLeuHisLeuGluLeuThrGluThrCysLeuAspMetMe 986
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 1266 GATT-----CCCCTTGAAATGA 1249
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 986 tAlaArgTyrValPheSerAsnPheThrAlaValProLysArgSerProValGlyGluPh 1006
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 1248 ACTTTTATTTGGTTTACTGACATTTATGTAGATTTCAGTGAAAGCTCTATAAATACA 1189
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 1006 eLeuLeuAlaGlyGlyArgThrLysThrTrpLeuValGlyAsnLysLeu----- 1022
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 1188 ATAATACGGGTTGAAAAGGCAGACATCTTAGTTGCATATATTACAGCTTTATCTTAC 1129
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 1023 -----ValThrVa 1025
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 1128 GGTCAGGCCATTGGAACTGCAATGTGGAGACTGTTTGTATATCAGACATGGAGAGGTGC 1069
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 1025 lThrThrSerValGlyThrGlyThrArgSerLeuLeuGlyLeuAspSerGlyGlu---Le 1044
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 1068 ACGTCTTAAAGGGGAGACAGCTGCTTCGGTTGGGAATCAT-----CACACTCCTCCGC 1015
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 1044 uGlnSerGlyProGluSerSerSerProGlyValHisValArgGlnThrLysGluAl 1064
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 1014 TCAGCGCGCTTCCCTTCCCGCTGTTTCACAGCTGCTTCCAGAGTTTGTCCAGCAA 955
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 1064 aProAlaValLeu-LeuSerGlnAla---GlyGlnGlnValSerArgGlyAlaArgAspA 1083
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 954 GGAATAAATGAATGCATACAGGA-----CTTTGGCTA 922
||| : : : : : ||| ||| ||| : : : ||| : : :
Db 1083 rgValArgSerMetSerGlyGlyHisGlyLeuArgValGlyAlaLeuAspValProAlas 1103
||| : : : : : ||| ||| ||| : : : ||| : : :
Qy 921 GTAGACTGTCTGGGTATTGTGAGCATCAGGT-----TGTTGTTCTCTGTTATTATAAT 869
||| : : : : : ||| ||| ||| : : : ||| : : :
```

```
Db      1103 erGlnPheLeuGlySerAlaThrSerProGlyProArgThrAlaProAlaAlaLysProG 1123
Qy      867 AAAGTCAACAGTGAAGTTCACACATTTCCATCTCTGTGTATACACCAAGTCAAACTTC- 809
Db      1123 lulyAlaSerAlaGlyThrArgValProVal-----GlnGluLysThrAsnLeuA 1140
Qy      808 --TCATTATGAGACAGGTGAAGATGTCCTTCCACAGCCCAACTCTTCCCGGCCACGG 751
Db      1140 laAlaTyValProLeuLeuThrGlnGlyTrpAlaGluLeuValArgArgProThrG 1160
Qy      750 GCAGGGCTCAGCGTTG----- 734
Db      1160 lyAsnThrSerTrpLeuMetSerLeuGluAsnProLeuSerProPheSerSerAspIleA 1180
Qy      733 -----CAGC 730
Db      1180 enAsnMetProLeuGlnGluLeuSerAsnAlaLeuMetAlaAlaGluArgPheLysGluH 1200
Qy      729 ACTGGTCCACAGCTGGTGGCAG--TGCTCCGTCAGGGCTTCAGCTGGGCTTGTCTAC 673
Db      1200 isArgAspThrAlaLeuTyrlsSerLeuSerValProAlaAlaSerThrAlaLysProp 1220
Qy      672 AGGACACCCCGAGAAAT-----CTGGCCAGCTGCTCCACCATCGTCACAGCTCCCGAT 619
Db      1220 roProLeuProArgSerAsnThrValAlaSerPheSerSerLeuTyx---GlnSerSerC 1239
Qy      618 GCATGTCTTCATACTTGAGAAAAGCAGCTTC----- 587
Db      1239 ysGlnGlyGlnLeuHisArgSerValSerTrpAlaAspSerAlaValMetGluGluG 1259
Qy      586 -----GAGTCATCCGCTGCTCCAGAACTCTCCAGAACTCTCTGCA 556
Db      1259 lySerProGlyGluValProValLeuValGlu-ProproGlyLeuGluAspValGluAla 1278
Qy      555 CGTGCTCAAAACGAGCGCGTAGCCAGCTTATCATTCATAAACTCC----- 508
Db      1279 AlaLeuGlyMetAspArgArgThrAspAlaTyrlsSerArgSerSerValSerSerGln 1298
Qy      507 -----GCCAAGTCTTTGAAGGTGC 487
Db      1299 GluGluLysSerLeuHisAlaGluLeuValGlyArgGlyIleProIleGluArgVal 1318
Qy      486 CTCGGTAGCTCATGTCGCGCAGAGCGGTGGAACTGATATAGACACCAACCATCTCT 427
Db      1319 ValSerSer-GluGlyGlyArgProSerValAspLeuSerPheGlnProSerGlnProLe 1338
Qy      426 TGGGGTTCCGAGCCATATAGATACCTTTGGAGTCTCCATTGTGGAGGTTCAGAGGGCAGAA 367
Db      1338 uSerLysSerSerSerProGluLeu--GlnThrLeuGlnAspIleLeuGlyAspPr 1357
Qy      366 AGCGGTAGGGCAGGTGGCTCTTGATAGAGCGGGGAGAGGTGATGCTCTTGATGATGCCA 307
Db      1357 oGly-----AspLysAlaAspValGlyArgLeuSerProGluVally 1371
Qy      306 GGCCCGGCT---GTGGTACTCCAGGACCGGAGCTGCTCGTATGTTTCATCAAGCCG 251
Db      1371 sAlaArgSerGlnSerGlyThrLeuAspGlyGluSerAlaAlaTrpSerAlaSerGlyG 1391
Qy      250 ATCTCATCGGGGTGAGCGCCCTGGGTCA---CCAAGTAGACCACTCTCTGCAAGAGCTG 194
Db      1391 uAspSerArgGlyGlnProGluGlyProLeuProSerSerSerProArgSerProSerGl 1411
Qy      193 GTGCCGAGTGGGGTAGGTACGATCCACAGCTGCTGGGCGGACCGCCACGCCATCGAAGTAC 134
Db      1411 yLeuArgProArgGlyTyThrIleSerAsp----- 1421
Qy      133 ATCTCCTCCATCTTCCCGCGCAGAGGGCGGCGGACCGCCACGCCATCGAAGTAC 74
Db      1422 -SerAlaProSerArgArgGlyLysArgValGluArgAspAlaLeuLysSerArgAlaTh 1441
Qy      73 TTGCTCTCGAACTCCC-----CGGGG 52
```

```
Db      1441 rAlaSerAsnAlaGluLysValProGly 1450
RESULT 100
US-10-196-749-224
; Sequence 224, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; PRIOR FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 224
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-224
```

```
Alignment Scores:
Pred. No.:      8.36      Length:      341
Score:          81.00     Matches:     51
Percent Similarity: 33.7%   Conservative: 17
Best Local Similarity: 25.2% Mismatches:    61
Query Match:     1.8%     Indels:      74
DB:              6       Gaps:       13
```

```
US-10-768-158-1 (1-2419) x US-10-196-749-224 (1-341)
Qy      594 GCAGTTCGAGTCCATCGGTGCTCCAGAACTCCTCCACGTGCTCAACCCAGG----- 541
Db      121 SerArgGluLeuProSerAlaThrProAsnThrAlaGlySerSerThrArgPheIle 140
Qy      540 -----AGCGGTAGCCAGCTTATCATTTCAATAACCTCC-----GGCAGA 502
Db      141 AlaAsnSerGlnGluProGluIleArgLeuThrSerSerLeuProArgSerProGlyArg 160
Qy      501 ATTCTTGAAGGTGCCTCGGT-----AGCTCATGCTCCGACGAGCGGT 457
Db      161 SerThrGluAspLeuProGlySerGlnAlaThrLeuSerGlnTrpSerThr---ProGly 179
```





**THIS PAGE BLANK (USPTO)**



85 289.5 6.6 312 4 US-10-072-012-665 Sequence 665, App  
86 281.5 6.4 303 4 US-10-072-012-250 Sequence 250, App  
87 281 6.4 317 4 US-10-437-963-157482 Sequence 157482,  
88 277 6.3 345 6 US-11-096-568A-1016 Sequence 1016, Ap  
89 275.5 6.3 320 3 US-09-854-122-20 Sequence 20, App1  
90 275.5 6.3 320 6 US-11-033-030-20 Sequence 20, App1  
91 275 6.2 340 6 US-11-096-568A-12566 Sequence 12566, A  
92 274 6.2 282 6 US-11-096-568A-1018 Sequence 1018, Ap  
93 274 6.2 286 6 US-11-096-568A-1017 Sequence 1017, Ap  
94 269.5 6.1 336 4 US-10-437-963-198983 Sequence 198983,  
95 268.5 6.1 329 6 US-11-096-568A-7712 Sequence 7712, Ap  
96 268.5 6.1 330 6 US-11-096-568A-7711 Sequence 7711, Ap  
97 268.5 6.1 340 6 US-11-096-568A-7710 Sequence 7710, Ap  
98 267.5 6.1 344 4 US-10-767-701-42402 Sequence 42402, A  
99 267 6.1 343 4 US-10-259-194A-330 Sequence 330, App  
100 267 6.1 343 4 US-10-437-963-159374 Sequence 159374,

## ALIGNMENTS

## RESULT 1

US-10-427-631-7  
; Sequence 7, Application US/10427631  
; Publication No. US20030175923A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE CORPORATION; TANG, Y. Tom;  
; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;  
; APPLICANT: BAUGHN, Mariah R.; LAL, Preeti G.;  
; APPLICANT: YUE, Henry; HILLMAN, Jennifer L.;  
; APPLICANT: AZIMZAI, Yalda  
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS  
; FILE REFERENCE: PF-0592-1 DIV  
; CURRENT APPLICATION NUMBER: US/10/427,631  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: US 09/786,240  
; PRIOR FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: PCT/US99/20989  
; PRIOR FILING DATE: 1999-09-09  
; PRIOR APPLICATION NUMBER: US 60/172,220  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: US 60/155,248  
; PRIOR FILING DATE: 1998-11-04  
; PRIOR APPLICATION NUMBER: US 60/133,642  
; PRIOR FILING DATE: 1999-05-11  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PERL Program  
; SEQ ID NO 7  
; LENGTH: 284  
; TYPE: PPT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030175923A1 1784742CD1  
US-10-427-631-7

Alignment Scores:  
Pred. No.: 1.71e-119 Length: 284  
Score: 1530.00 Matches: 284  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 34.8% Indels: 0  
Db: 4 Gaps: 0

US-10-768-158-1 (1-2419) x US-10-427-631-7 (1-284)

QY 21 ATGGGGAGAGCGAGCGCCGAGACCCCGGGAGTTCGAGAGCAAGTACTTC 80  
|||||  
Db 1 MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerIstYrPhe 20  
QY 81 GAGTTCCATGGCGTGGCTCCGCCCTTCTGCGCGGGAAGATGGAGGATCGCCAAC 140  
|||||  
Db 21 GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluGluIleAlaAsn 40

QY 141 TTCCCGTGGCGCCAGCGAGCTGTGGATCGTCCACCTACCCCAAGTCCGGCACCAGCTTG 200  
|||||  
Db 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60  
QY 201 CTGAGGAGGTGGTCTACTTGGTGAGCGAGCGCGCTGACCCCGATGAGATGGCTTGATG 260  
|||||  
Db 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80  
QY 261 AACATCAGCAGCAGCTCCCGGCTCGGAGTACCCAGCGCGGCGCTGGACATCATCAAG 320  
|||||  
Db 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100  
QY 321 GAATGACCTCTCCCGCGCTCATCAAGAGCCACCTGACCGCTTCTGCGCTCTGAC 380  
|||||  
Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120  
QY 381 CTCCACATCGAGACTCCAAGGTCTATATGGCTCGCAACCCCAAGGATCTCGTGGTG 440  
|||||  
Db 121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuValVal 140  
QY 441 TCTTATTATCAGTCCACCGCTCTCTCGGACCATGAGCTACCGAGGACCTTTCAAGAA 500  
|||||  
Db 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160  
QY 501 TTCTGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCTCGTGGTTTGACACGTGCG 560  
|||||  
Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180  
QY 561 GAGTTCTGGGAGCACCCTGAGCTCGAAGCTGCTTTTCTCAAGTATGAAGACATGCAT 620  
|||||  
Db 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHis 200  
QY 621 CGGACCTGTGAGCATGGTGGAGCAGCTGCCAGATTCCTGGGGGTGCTGTGACAAG 680  
|||||  
Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220  
QY 681 GCCCAGCTGGAAGCCCTGACGAGCAGCTGCCACAGCTGGTGGACAGTGTGCTCAACGCT 740  
|||||  
Db 221 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysAsnAla 240  
QY 741 GAGCCCTGCGCGTGGCGCGGGAAGATGGGCTGTGGAGGACATCTTCACCGTCTCC 800  
|||||  
Db 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 260  
QY 801 ATGAATCAGAGTTGACTTGGTGTATAAACAGAGATGGAAAGTGTGACCTCACGTTT 860  
|||||  
Db 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280  
QY 861 GACTTTTATTTA 872  
|||||  
Db 281 AspPheTyrLeu 284

## RESULT 2

US-10-757-262-110  
; Sequence 110, Application US/10757262  
; Publication No. US20040197825A1  
; GENERAL INFORMATION:  
; APPLICANT: Karicheti, Venkateswarlu  
; APPLICANT: Silos-Santiago, Inmaculada  
; APPLICANT: Eliasof, Scott D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING  
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,  
; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,  
; TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43239, 44373, 51164,  
; TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,  
; TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 8675,  
; TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,  
; TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,  
; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR  
; TITLE OF INVENTION: 55053  
; FILE REFERENCE: MP103-007P1RNMNIM  
; CURRENT APPLICATION NUMBER: US/10/757,262  
; CURRENT FILING DATE: 2004-01-14

; PRIOR APPLICATION NUMBER: US 60/440,318
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/444,783
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/457,901
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/468,775
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/471,614
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/478,742
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/488,529
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/491,156
; PRIOR FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US 60/499,594
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/506,332
; PRIOR FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 110
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-757-262-110

Alignment Scores:
Pred. No.: 1,71e-119 Length: 284
Score: 1530.00 Matches: 284
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 34.8% Indels: 0
DB: 4 Gaps: 0

US-10-768-158-1 (1-2419) x US-10-757-262-110 (1-284)

Qy 21 ATGGCGGAGAGCGAGCGGAGACCCCGGGGAGTTCGAGAGCAAGTACTTC 80
Db 1 MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerLysTyrPhe 20
Qy 81 GAGTTCCATGCGTCCGCTCCGCTTCCTCCGCGGGAAGATGAGAGATCGCCAAC 140
Db 21 GluPheHisGlyValArgLeuProProPheCysArgGlyLysMetGluGluLeuAlaAsn 40
Qy 141 TTCGGTGGCGGCGGAGCGAGTGTGGATGTCACCTACCCCAAGTCCGGCAGCAGCTTG 200
Db 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60
Qy 201 CTGCAGGAGGTGGTCTACTTGTGAGCAGGCGCTGACCCCGATGAGATCGGCTTGATG 260
Db 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
Qy 261 AACATCGACGAGCTCCCGTCTCTGGAGTACCCACAGCCGCGCTCGACATCATCAAG 320
Db 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100
Qy 321 GAATGACTCTCCCGCTCATCAAGAGCCACTGCTCCGCTACCGCTTCTCCCTCTGAC 380
Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120
Qy 381 CTCACAATGAGACTCCAAAGTCTATATGCTCGCAACCCCAAGGATCTGGTGTG 440
Db 121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuVal 140
Qy 441 TCTTATTACAGTTCACCGCTCTCTGGGAGCATGAGCTACCGAGGACCTTTCAAGAA 500
Db 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160
Qy 501 TTCTCCCGAGGTTTATGAATGATAGCTGGGCTACGGCTCCTGGTTTGGAGCAGTGCAG 560
Db 161 PheCysArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180

Qy 561 GAGTTCTGGGAGCACCCGATGGACTCGAACGTGCTTTTCTCAAGTATGAAGACATGCAT 620
Db 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHis 200
Qy 621 CGGACCTGGTGGAGATGGTGGACAGCTGCCAGATTCCTGGGGGTCTCTGTGACAAG 680
Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220
Qy 681 GCCCAGCTGGAAGCCCTGACGGAGCACTGCCACAGCTGGTGGACCAAGTGTGCAACGCT 740
Db 221 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240
Qy 741 GAGCCCTGCCGCTGGCGCGGGAAGAGTTGGGTGTGGAGGACATCTTCACCGTCTCC 800
Db 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 260
Qy 801 ATGATCAGAGAGTTGACTTGGTCTATATAACAGAGATGGAAAGTGTGACCTCACGTTT 860
Db 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280
Qy 861 GACTTTTATTATA 872
Db 281 AspPheTyrLeu 284

RESULT 3

US-10-768-158-2
; Sequence 2, Application US/10768158
; Publication No. US20040204359A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Karicheti, Venkateswarlu
; APPLICANT: Eliasof, Scott D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 16386, 15402, 21165, 1423,
; TITLE OF INVENTION: 636, 12303, 21425, 27410, 38554, 38555, 55063, 57145, 59914,
; TITLE OF INVENTION: 94921, 16852, 33260, 58573, 30911, 85913, 14303, 16816,
; TITLE OF INVENTION: 17827 OR 32620
; FILE REFERENCE: MPI03-012P1RNMNMIM
; CURRENT APPLICATION NUMBER: US/10768,158
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: US 60/444,781
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/452,291
; PRIOR FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: US 60/454,540
; PRIOR FILING DATE: 2003-03-13
; PRIOR APPLICATION NUMBER: US 60/478,805
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/491,048
; PRIOR FILING DATE: 2003-07-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-768-158-2

Alignment Scores:
Pred. No.: 1,71e-119 Length: 284
Score: 1530.00 Matches: 284
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 34.8% Indels: 0
DB: 4 Gaps: 0

US-10-768-158-1 (1-2419) x US-10-768-158-2 (1-284)

Qy 21 ATGGCGGAGAGCGAGCGGAGACCCCGGGGAGTTCGAGAGCAAGTACTTC 80
Db 1 MetAlaGluSerGluAlaGluThrProSerThrProGlyGluPheGluSerLysTyrPhe 20

```
Qy 81 GAGTTCCATGGCGTGGCGCTGCCCGCTTCTGCGCGGAGAGATGGAGATCGCCAAAC 140
Db 21 GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluGluValAlaAsn 40
Qy 141 TTCCCGGTGGCGGCCAGCGAGTGTGGATCTCACTACCCCAAGTCCGGCACCAGCTTG 200
Db 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60
Qy 201 CTGACGAGGTGTCTACTTGGTGAGCGAGCGGCGCTGACCCCGATGAGATCGGCTTGATG 260
Db 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
Qy 261 AACATCCAGCAGCAGCTCCCGTCTCGAGTACCACAGCGCGCTGGACATCATCAAG 320
Db 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100
Qy 321 GAATGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTTCTGCCCTCTGAC 380
Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120
Qy 381 CTCACAAATGAGACTCAAGGTCAATATATGCTGCGAATCCCAAGGATCTGGTGGTG 440
Db 121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuVal 140
Qy 441 TCTTATTATCAGTTCCACCGCTCTCGGACCATGAGCTACCGAGGACACCTTTCAAGAA 500
Db 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160
Qy 501 TTCTGCCGAGGTTTATGAATGATAAGCTGGGCTACGGCTCTCTGGTTTGACACGTGCAG 560
Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180
Qy 561 GAGTTCTGGAGCACCAGCATCGAATGATGATGATGATGATGATGATGATGATGATGATG 620
Db 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysTyrGluAspMetHis 200
Qy 621 CGGACCTGGTGAGCAGTGGTGAGCAGCTGCCAGCTGCTGGGCGGTCTCTGTGCAAG 680
Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220
Qy 741 GCCAGCTGGAAGCCCTGACGAGCAGCTGCCACAGCTGGTGAGACATCTTCAACGCT 740
Db 241 GluAlaLeuProValGlyArgGlyValGlyLeuTrpLysAspIlePheThrValSer 260
Qy 801 ATGAATGAGAAGTTGACTTGGTGATATAACAGAGATGGGAAAGTGTGACCTCACGTTT 860
Db 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280
Qy 861 GACTTTTATTTA 872
Db 281 AspPheTyrLeu 284
```

## RESULT 4

```
US-10-205-331-46
; Sequence 46, Application US/10205331
; Publication No. US20040058326A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnoch, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018199
; CURRENT APPLICATION NUMBER: US/10/205,331
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
```

```
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Sulfotransferase-like protein
US-10-205-331-46
```

```
Alignment Scores:
Pred. No.: 1,79e-117 Length: 284
Score: 1506.00 Matches: 278
Percent Similarity: 99.3% Conservative: 4
Best Local Similarity: 97.9% Mismatches: 2
Query Match: 34.2% Indels: 0
DB: 4 Gaps: 0
```

US-10-768-158-1 (1-2419) x US-10-205-331-46 (1-284)

```
Qy 21 ATGCGGAGAGCGGCGGAGACCCAGCAGCAGCCCGGGGAGTTCGAGAGCAAGTACTTC 80
Db 1 MetAlaGluSerGluAlaGluThrProGlyThrProGlyGluPheGluSerLysTyrPhe 20
Qy 81 GAGTTCCATGGCGTGGCGCTTCTGCGCGGAGAGATGGAGAGATCGCCAAAC 140
Db 21 GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluAspIleAlaAsp 40
Qy 141 TTCCCGGTGGCGGCCAGCGAGCTGTGGATCGTCACCTACCCCAAGTCCGGCACCAGCTTG 200
Db 41 PheProValArgProSerAspValTrpIleValThrTyrProLysSerGlyThrSerLeu 60
Qy 201 CTCGAGCAGTGTCTACTTGGTGAGCCAGCGCGCTGACCCCGATGAGATCGGCTTGATG 260
Db 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProAspGluIleGlyLeuMet 80
Qy 261 AACATCCAGCAGCAGCTCCCGCTCTCGAGTACCACAGCGGCGCTTGGACATCATCAAG 320
Db 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100
Qy 321 GAATGACCTCTCCCGCTCATCAAGAGCAGCTGCCCTTACCGCTTTCTGCCCTCTGAC 380
Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120
Qy 381 CTCACAAATGAGACTCCAAAGTTCATCTATATGCTCGCAACCCCAAGGATCTGCTGGTG 440
Db 121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuVal 140
Qy 441 TCTTATTATCAGTTCCACCGCTCTCTCGGAGCAGCTACCGAGCAGCTTTTCAAGAA 500
Db 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160
Qy 501 TTCTGCCGAGGTTTATGAATGATAGCTGGGTACGGCTCTCGGCTCTGCTGGTGTGACGCTG 560
Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180
Qy 561 GAGTTCTGGAGCAGCCGATGGATCGAACGTGTTTTTCTCAAGTATGAAGACATGAT 620
Db 181 GluPheTrpGluHisArgMetAspAlaAsnValLeuPheLeuValAspGlnCysCysAsnAla 240
Qy 621 CGGACCTGGTGAGTGGTGAGCAGCTGCCACAGATTCCTGGGGGTGTCTGTGCAAG 680
Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220
Qy 681 GCCAGCTGGAAGCCCTGAGCGAGCAGCTGCCACAGCTGGTGGACCAAGTGTGCAACGCT 740
Db 221 AlaGlnLeuGluSerLeuIleGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240
Qy 741 GAGCGCTGCGCGCGGCGGAGAGTGGGTGTGGAGGAGCATCTTCAACGCTCTCC 800
Db 241 GluAlaLeuProValGlyArgGlyValGlyLeuTrpLysAspIlePheThrValSer 260
Qy 801 ATGAATGAGAAGTTTGAATTTGGTGATATAACAGAGATGGGAAAGTGTGACCTCACGTTT 860
```

```
Db      261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280
Oy      861 GACTTTTATTTA 872
Db      281 AspPheTyrLeu 284

RESULT 5
US-09-989-442-113
; Sequence 113, Application US/09989442
; Publication No. US20030013649A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0208
; CURRENT APPLICATION NUMBER: US/09/989,442
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
```

```

; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,065
; PRIOR FILING DATE: 2000-09-14

Alignment Scores:
Pred. No.: 2,5e-35 Length: 105
Score: 528.00 Matches: 101
Percent Similarity: 95.3% Conservative: 0
Best Local Similarity: 95.3% Mismatches: 4
Query Match: 12.0% Indels: 1
DB: 3 Gaps: 0

US-10-768-158-1 (1-2419) x US-09-989-442-113 (1-105)

```

380 CCTCCAAATGAGACTCCAGCTCATCTATATGGCTCGCAACCCCAAGGATCTGGTGGT 439  
 1 ProProGlnTrpArgLeuGlnGlyHisLeuTyGlySerGlnProGlnGlySerGlyGly 20  
 440 GTCTTATTATCAGTTCACCGCTCTCTGCGGACCATGAGCTACCGAGGCACCTTTCAAGA 499  
 21 ValLeuLeuSerValProProLeuSerAlaAspHisGluLeuProArg\*\*\*LeuSerArg 40  
 500 ATTCTCCGAGGTTTATGAATAGCTGGGCTACGGCTCTCTGGTTTGAGCACGTGCA 559  
 41 IleLeuProGlu-PheMetAsnAspLysLeuGlyTyGlySerTrpPheGluHisValGI 60  
 560 GGAGTTCTGGGAGCACCGCATGGACTCGAACGTCCTTTTCTCAAGTATGAACATGCA 619  
 60 nGluPheTrpGluHisArgMetAspSerAenValLeuPheLeuLysTyGluAspMetHi 80  
 620 TCGGGACCTGGTACCATGCTGGAGCAGCTGGCCAGATTCTGGGGGTGCTCTGTGCAA 679  
 80 sArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCys\*\*\*I 100  
 680 GGCCGAGCTGGAAGCC 695  
 100 ePheGlnLeuGluAla 105

RESULT 6  
 US-10-072-012-667  
 ; Sequence 667, Application US/10072012  
 ; Publication No. US20040033493A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tchernev, Velizar  
 ; APPLICANT: Spytek, Kimberly  
 ; APPLICANT: Zernhusen, Bryan  
 ; APPLICANT: Patturajan, Meera  
 ; APPLICANT: Shimkets, Richard  
 ; APPLICANT: Li, Li  
 ; APPLICANT: Gangolli, Esha  
 ; APPLICANT: Padigar, Muralidhara  
 ; APPLICANT: Anderson, David W.  
 ; APPLICANT: Rastelli, Luca  
 ; APPLICANT: Miller, Charles E.  
 ; APPLICANT: Gerlach, Valerie  
 ; APPLICANT: Taupier Jr, Raymond J.  
 ; APPLICANT: Gusev, Vladimir Y.  
 ; APPLICANT: Colman, Steven D.  
 ; APPLICANT: Wolenc, Adam R.  
 ; APPLICANT: Pena, Carol E. A  
 ; APPLICANT: Furtak, Katarzyna  
 ; APPLICANT: Grosse, William M.  
 ; APPLICANT: Alsobrook II, John P.  
 ; APPLICANT: Lepley, Denise M.  
 ; APPLICANT: Rieger, Daniel K.  
 ; APPLICANT: Burgess, Catherine E.  
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
 ; FILE REFERENCE: 21402-258  
 ; CURRENT APPLICATION NUMBER: US/10/072,012  
 ; CURRENT FILING DATE: 2002-01-31  
 ; PRIOR APPLICATION NUMBER: 60/265,102  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: 60/265,514  
 ; PRIOR FILING DATE: 2001-01-31  
 ; PRIOR APPLICATION NUMBER: 60/265,517  
 ; PRIOR FILING DATE: 2001-01-31  
 ; PRIOR APPLICATION NUMBER: 60/265,412  
 ; PRIOR FILING DATE: 2001-01-31  
 ; PRIOR APPLICATION NUMBER: 60/265,395  
 ; PRIOR FILING DATE: 2001-01-31  
 ; PRIOR APPLICATION NUMBER: 60/266,406  
 ; PRIOR FILING DATE: 2001-02-02  
 ; PRIOR APPLICATION NUMBER: 60/266,767  
 ; PRIOR FILING DATE: 2001-02-05  
 ; PRIOR APPLICATION NUMBER: 60/267,057  
 ; PRIOR FILING DATE: 2001-02-07





345	LySerGlyThrThrTrpValSerGlnIleLeuAspMetIleTyrGlnGlyGlyAspLeu	364
243	GATGAGATCGGCTTGATGAACATCATGACGAGCAGCTCCGGCTCCTGGAGTACCCACAGCGG	302
365	GluLysCysAsnArgAlaProIleTyrValArgValProPheLeuGluValAsnAspPro	384
303	-----GGCCTGGACATCATCAAGNAACGTGACCTCTCCCGGCTCATCAAGAGC	350
385	GlyGluProSerGlyLeuGluThrLeuLysAspThrProProArgLeuIleLysSer	404
351	CACCTGCCCTACCGCTTCTGCCCTCTACCTCCACAATGGAGACTCCAAAGTGCATCAT	410
405	HisLeuProLeuAlaLeuLeuProGlnThrLeuLeuAspGlnLysValLysValTyr	424
411	ATGGCTGCAACCCCAAGGATCTGGTGTGTCTTATTATCATCTCCACCGCTCTCTGGG	470
425	ValAlaArgAsnProLysAspValAlaValSerTyrTyrHisPheHisArgMetGluLys	444
471	ACCATGAGCTACCGAGGCACCTTTCAAGAATTTCTCCGGAGGTTTATGAATGATAAGCT	530
445	AlaHisProGluProGlyThrTrpAspSerPheLeuGluLysPheMetAlaGlyGluVal	464
531	GGCTACGGCTCCTGGTTTGACGACGTGACGAGGTCTCGGAGCACCGCATGACTCGCAAC	590
465	SerTyrGlySerTrpTyrGlnHisValGlnGluTrpTrpGluLeuSerArgThrHisPro	484
591	GTGCTTTTTCCTCAAGATGAGACATGCATCGGACCTGGTGACGATGGTGGAGCAGCTG	650
485	ValLeuTyrLeuPheTyrGluAspMetLysGluAsnProLysArgGluIleGlnLysIle	504
651	GCCAGATTCTCGGGGCTCTCTGTGCACAGGCCACAGCTGGAAAGCCCTCAGCGACACTGC	710
505	LeuGluPheValGlyArgSerLeuProGluGluThrMetAspPheMetValGlnHisThr	524
711	-----CACCAGCTG	719
525	SerPheLysGluThrLysLysAsnProMetThrAsnTyrThrThrValProGlnGluLeu	544
720	GTGGACCAGTCTGCAACGCTGAGGCCCTGCCC---GTGGCGCGGGAGAGATTGGGCTG	776
545	MetAspHisSerIleSer-----ProPheMetArgLysGlyMetAlaGlyAsp	560
777	TGGAAAGCACATCTTCACCGCTCTCCATGAATGAGAAGTTTGACTGGTGTATATAACAGAG	836
561	TrpLysThrThrPheThrValAlaGlnAsnGluArgPheAspAlaAspTyrAlaGluLys	580
837	ATGGGAAGTGTGACCTCACGTTT	860
581	MetAlaGlyCysSerLeuSerPhe	588

RESULT 8  
 US-09-854-122-21  
   ; Sequence 21, Application US/09854122  
   ; Patent No. US20020016980A1  
   ; GENERAL INFORMATION:  
   ; APPLICANT: ALBERTE, RANDALL S.  
   ; APPLICANT: SMITH, ROBERT  
   ; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA  
   ; FILE REFERENCE: PHA-007.01  
   ; CURRENT APPLICATION NUMBER: US/09/854,122  
   ; CURRENT FILING DATE: 2001-09-10  
   ; PRIOR APPLICATION NUMBER: 60/202,529  
   ; PRIOR FILING DATE: 2000-05-10  
   ; NUMBER OF SEQ ID NOS: 51  
   ; SOFTWARE: PatentIn Ver. 2.1  
   ; SEQ ID NO 21  
   ; LENGTH: 295  
   ; TYPE: FRT  
   ; ORGANISM: Homo sapiens  
 US-09-854-122-21  
 Alignment Scores:  
 Pred. No.: 1.9e-30 Length: 295

```
; CURRENT APPLICATION NUMBER: US/10/769,507
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: 10/024,628
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: 09/288,037
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/081,037
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-769-507-4

Alignment Scores:
Pred. No.: 1.9e-30      Length: 295
Score: 472.00          Matches: 100
Percent Similarity: 54.5%      Conservative: 52
Best Local Similarity: 35.8%    Mismatches: 101
Query Match: 10.7%            Indels: 26
DB: 5                      Gaps: 4

US-10-768-158-1 (1-2419) x US-10-769-507-4 (1-295)
Qy 90 GGGGTGGCGTCCCGCCCTTCTGCGCGGGAAGATGGAGGAGATCGCAACTTCCCGGTG 149
Db 17 GlyValProLeuIleTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
Qy 150 CGGCCAGCAGCGTGTGGATCGTACCTACCCCAAGTCCCGCACCAGCTGCTCCAGGAG 209
Db 37 ArgProAspLeuLeuIleAsnThrTyrProLysSerGlyThrTrpValSerGln 56
Qy 210 GTGTCTACTTGGTGAGCGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
Db 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysAsnArgAlaProIleTyr 76
Qy 270 GAGCAGCTCCCGTCTCGGAGTACCCACAGCCG-----GGCCTGGACATCATC 317
Db 77 ValArgValProPheLeuGluValAsnAspProGlyGluProSerGlyLeuGluThrLeu 96
Qy 318 AAGGAAGTACCTCTCCCGCTCATCAAGACGACCTGCCCTACCGCTTCTGCCCTCT 377
Db 97 LysAspThrProProArgLeuIleLeuSerHisLeuProLeuAlaLeuLeuProGln 116
Qy 378 GAGCTCCACAATGAGGACTCCAAAGTCACTATATGCTCGCAACCCCAAGGATCTGGT 437
Db 117 ThrLeuLeuAspGlnLysValLysValValValValValValValValValValVal 136
Qy 438 GTGTCTTATTATCAGTTCCAGCCCTCTCTCCGAGACCATGAGCTACCGAGGACCTTCAA 497
Db 137 ValSerTyrTyrHisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAsp 156
Qy 498 GAATTCCTCCGGAGGTTTATGAATGATGATGATGATGATGATGATGATGATGATGATG 557
Db 157 SerPheLeuGluLysPheMetAlaGlyLysValSerTyrGlySerTyrGlnHisVal 176
Qy 558 CAGGAGTCTGGGAGCACCAGCATGAGTCCGAAGCTGCTTTTCTCAAGTATGAAGACATG 617
Db 177 GlnGluThrProGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
Qy 618 CATCGGACCTGGTGGACATGGTGGAGCTGCGCCAGATTCCTGGGGGGTGTCTCTGTGAC 677
Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
Qy 678 AAGGCCAGCTGGAGCCCTGACGGAGCATGCG-----GGCCTGGACATCATC 710
Db 217 GluGluThrMetAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236
Qy 711 -----CACCAGCTGGTGGACGACGATGCTCTCAACGCTGAGGCC 746
Db 237 MetThrAsnTyrThrThrValProGlnGluLeuMetAspHisSerIleSer----- 253

; CURRENT APPLICATION NUMBER: US/10/769,507
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: 10/024,628
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: 09/288,037
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/081,037
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-769-507-4

Alignment Scores:
Pred. No.: 1.9e-30      Length: 295
Score: 472.00          Matches: 100
Percent Similarity: 54.5%      Conservative: 52
Best Local Similarity: 35.8%    Mismatches: 101
Query Match: 10.7%            Indels: 26
DB: 5                      Gaps: 4

US-10-768-158-1 (1-2419) x US-11-033-030-21 (1-295)
Qy 90 GGGGTGGCGTCCCGCCCTTCTGCGCGGGAAGATGGAGGAGATCGCAACTTCCCGGTG 149
Db 17 GlyValProLeuIleTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
Qy 150 CGGCCAGCAGCGTGTGGATCGTACCTACCCCAAGTCCCGCACCAGCTGCTCCAGGAG 209
Db 37 ArgProAspLeuLeuIleAsnThrTyrProLysSerGlyThrTrpValSerGln 56
Qy 210 GTGTCTACTTGGTGAGCGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
Db 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysAsnArgAlaProIleTyr 76
Qy 270 GAGCAGCTCCCGTCTCGGAGTACCCACAGCCG-----GGCCTGGACATCATC 317
Db 77 ValArgValProPheLeuGluValAsnAspProGlyGluProSerGlyLeuGluThrLeu 96
Qy 318 AAGGAAGTACCTCTCCCGCTCATCAAGACGACCTGCCCTACCGCTTCTGCCCTCT 377
Db 97 LysAspThrProProArgLeuIleLeuSerHisLeuProLeuAlaLeuLeuProGln 116
Qy 378 GAGCTCCACAATGAGGACTCCAAAGTCACTATATGCTCGCAACCCCAAGGATCTGGT 437
Db 117 ThrLeuLeuAspGlnLysValLysValValValValValValValValValValVal 136
Qy 438 GTGTCTTATTATCAGTTCCAGCCCTCTCTCCGAGACCATGAGCTACCGAGGACCTTCAA 497
Db 137 ValSerTyrTyrHisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAsp 156
Qy 498 GAATTCCTCCGGAGGTTTATGAATGATGATGATGATGATGATGATGATGATGATGATG 557
Db 157 SerPheLeuGluLysPheMetAlaGlyLysValSerTyrGlySerTyrGlnHisVal 176
Qy 558 CAGGAGTCTGGGAGCACCAGCATGAGTCCGAAGCTGCTTTTCTCAAGTATGAAGACATG 617
Db 177 GlnGluThrProGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
Qy 618 CATCGGACCTGGTGGACATGGTGGAGCTGCGCCAGATTCCTGGGGGGTGTCTCTGTGAC 677
Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
Qy 678 AAGGCCAGCTGGAGCCCTGACGGAGCATGCG-----GGCCTGGACATCATC 710
Db 217 GluGluThrMetAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236
Qy 711 -----CACCAGCTGGTGGACGACGATGCTCTCAACGCTGAGGCC 746
Db 237 MetThrAsnTyrThrThrValProGlnGluLeuMetAspHisSerIleSer----- 253
```

QY 558 CAGAGTTCTGGAGCACCAGCATGCGTCTCAAGTATGAAGCATG 617  
Db 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196  
QY 618 CATCGGACCTGGTGACGATGGTGAGCAGCTGCCAGATTCCTCGGGGGTCTCTGTGAC 677  
Db 197 LysGluAsnProLysArgGluLeuGlnLysLeuGluPheValGlyArgSerLeuPro 216  
QY 678 AAGCCCGCAGCTGGAAGCCCTGACGAGCATGTC----- 710  
Db 217 GluGluThrMetAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236  
QY 711 -----CACCAGCTGTGGACCGTCTGCAACGCTGAGGCC 746  
Db 237 MetThrAsnTyrThrValProGlnGluLeuMetAspHisSerIleSer----- 253  
QY 747 CTGCCC--GTGGCCCGGGAGAGTTGGGCTGTGGAGGACATCTTCACCGTCTCCATG 803  
Db 254 ---ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGln 272  
QY 804 AATGAGAAGTTGATGGTGATATAAAGAGATGGGAAGTGTGACCTCACGTTT 860  
Db 273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

## RESULT 11

US-10-072-012-666  
; Sequence 666, Application US/10072012  
; Publication No. US20040033493A1  
; GENERAL INFORMATION:  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Zernusen, Bryan  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Li, Li  
; APPLICANT: Gangolli, Esha  
; APPLICANT: Padigari, Muralidhara  
; APPLICANT: Anderson, David W.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Taupier Jr, Raymond J.  
; APPLICANT: Gusev, Vladimir Y.  
; APPLICANT: Colman, Steven D.  
; APPLICANT: Wolenc, Adam R.  
; APPLICANT: Pena, Carol E. A  
; APPLICANT: Furtak, Katarzyna  
; APPLICANT: Grosse, William M.  
; APPLICANT: Alsobrook II, John P.  
; APPLICANT: Lepley, Denise M.  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Burgess, Catherine E.  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-258  
; CURRENT APPLICATION NUMBER: US/10/072,012  
; CURRENT FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: 60/265,102  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 60/265,514  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,517  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,412  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,395  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/266,406  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 60/266,767  
; PRIOR FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: 60/267,057  
; PRIOR FILING DATE: 2001-02-07

; PRIOR APPLICATION NUMBER: 60/266,975  
; PRIOR FILING DATE: 2001-02-07  
; PRIOR APPLICATION NUMBER: 60/267,459  
; PRIOR FILING DATE: 2001-02-08  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1391  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 666  
; LENGTH: 296  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-072-012-666

Alignment Scores: 5.01e-30 Length: 296  
Pred. No.: 467.00 Matches: 100  
Score: 55.4% Conservative: 58  
Percent Similarity: 35.1% Mismatches: 105  
Best Local Similarity: 10.6% Indels: 22  
Query Match: 4 Gaps: 5  
DB:

US-10-768-158-1 (1-2419) x US-10-072-012-666 (1-296)  
QY 66 GAGAGCAAGTACTTCGAGTTCCATGGCGTGGCGCTTCTGCCGCGGAAGATG 125  
Db 10 GlnThrLysLeuLysGluValAlaGlyIleProLeuGlnAlaProThrValAsnTrp 29  
QY 126 GAGAGATCCCAACTTCCCGTGGCGGCCAGCGACGTGTGGATCGTACCTACCCCAAG 185  
Db 30 SerGlnIleGlnThrPheLysAlaLysProAspAspLeuLeuIleCysThrTyrProLys 49  
QY 186 TCCGGACACAGCTTGTGGAGAGGTGCTTACTTGTGTGAGCCAGGCGCTGACCCCGAT 245  
Db 50 SerGlyThrTrpIleGlnGluIleValAspMetIleGluInAsnGlyAspValGlu 69  
QY 246 GAGATCGGCTTGATGAACATCGACGAGCTCCCGTCTCTGGAGTAC-----CCA 296  
Db 70 LysCysGlnArgThrIleIleGlnHisArgHisProPheIleGluTrpAlaAArgProPro 89  
QY 297 CAGCCG--GGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCAC 353  
Db 90 GlnProSerGlyValAspLysAlaAsnAlaMetProAlaProArgIleLeuArgThrHis 109  
QY 354 CTGCCCTACGCTTCTGCGCTCTGACCTCCACAAATCGAGACTCCCAAGGTCTATATG 413  
Db 110 LeuProThrGlnLeuLeuProProSerPheTrpThrAsnAsnCysLysPheLeuTyrVal 129  
QY 414 GCTCGCAACCCCAAGGATCTGGTGTCTTATATCAGTTTCCACCGCTCTCTCGCGACC 473  
Db 130 AlaArgAsnAlaLysAspCysMetValSerTyrTyrHisPheTyrArgMetSerGlnVal 149  
QY 474 ATGAGCTACGAGCACCTTTCAGAAATTCGCGGAGGTTTATGATGATAGCTGGGC 533  
Db 150 LeuProAspProGlyThrTrpAsnGluTyrPheGluThrPheIleAsnGlyLysValSer 169  
QY 534 TAGCGCTCTGGTTGAGCAGCTGCGAGAGTTCTGGGACACCGCATGGACTCGAACTG 593  
Db 170 TrpGlySerTrpPheAspHisValLysGlyTrpTrpGluLeuArgAspArgTyrGlnIle 189  
QY 594 CTTTTCTCAAGTATGAAGACATGTCATCGGACCTGGTGTGACGATGGTGGAGCACCTGCC 653  
Db 190 LeuPheLeuPheTyrGluAspValLysArgAspProLysArgGluIleGlnLysValMet 209  
QY 654 AGATTCTGGGGGTGTC-----TGTCACAGGCCCGAGCTGGAGCCCTG 698  
Db 210 GlnPheMetGlyLysAsnLeuAspGluValValAspLysIleValLeuGluThrSer 229  
QY 699 ACGGAGCACTGCCACACAGCTGGTGTGCAACGTGTGACGCTGAGGCCCTGCC----- 752  
Db 230 PheGluLysMetLysGlu-----AsnProMetThrAsnArgSerThrValProLysSer 247  
QY 753 -----GTGGCGCGGGAAGAGTTGGCTGTGGAAGAC 785

```
Db 248 ValLeuaspGlnSerIleSerProPheMetArgLysGlyThrValGlyAspTrpLysAsn 267
Qy 786 ATCTTCACCGTCTCCATGAATGAGAAGTTTGACTTGGTGATATAACAGAAAGTGGAAAG 845
Db 268 HisPheThrValAlaGlnAsnAspArgPheAspGluIleTyrLysGlnLysMetGlyGly 287
Qy 846 TGTACCTCACGTTT 860
Db 288 ThrSerLeuAsnPhe 292

RESULT 12
US-10-206-839-3
; Sequence 1, Application US/10206839
; Publication No. US20030099977A1
; GENERAL INFORMATION:
; APPLICANT: Guida, Marco
; APPLICANT: Kurth, Janice
; TITLE OF INVENTION: Genotyping Human Phenol Sulfotransferase
; TITLE OF INVENTION: (STP2)
; FILE REFERENCE: 4389-6 (Formerly SEQ-16P)
; CURRENT APPLICATION NUMBER: US/10/206,839
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 09/328,174
; PRIOR FILING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 295
; TYPE: PRT
; ORGANISM: H. sapiens
US-10-206-839-3

Alignment Scores:
Pred. No.: 4.22e-29 Length: 295
Score: 456.00 Matches: 94
Percent Similarity: 54.5% Conservative: 58
Best Local Similarity: 33.7% Mismatches: 101
Query Match: 10.4% Indels: 26
DB: 4 Gaps: 4

US-10-768-158-1 (1-2419) x US-10-206-839-3 (1-295)
Qy 90 GCGGTGGCGTCCGCCCTTCTGCGCGGGAAGATGGAGAGATCGCCAACTTCCCGTG 149
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
Qy 150 CGGCCCGACGCTGTGGATCGTACCTACCCAGTCCGCGACCGCTTCTCGCAGGAG 209
Db 37 ArgProaspAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpValSerGln 56
Qy 210 GTGGTCTACTTGTGAGCAGCGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
Db 57 IleLeuaspMetIleTyrGlnGlyAspLeuGluLysCysHisArgAlaProIlePhe 76
Qy 270 GAGCAGTCCCGGTCTGGATGATACCCACGCGG-----GGCTGGACATCATC 317
Db 77 MetArgValProPheLeuGluPheLysValProGlyIleProSerGlyMetGluThrLeu 96
Qy 318 AAGNACTGACTCTCCCGGCTCATCAAGAGCCACCTGACCCCTACCGCTTCTGCCCTCT 377
Db 97 LysAsnThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116
Qy 378 GACCTCCACAATGGAGACTCCAAAGTCTATATGATGGCTCGCAACCCCAAGATCTGGTG 437
Db 117 ThrLeuLeuAspGlnLysValValValTyrValAlaAlaArgAsnAlaLysAspValAla 136
Qy 438 GTGCTTATTATCATGTTCCACCGTCTCTCGCGGACCATGAGCTACCGAGCACCTTCAA 497
Db 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValTyrProHisProGlyThrTrpGlu 156
Qy 498 GAATTCCTCCGGAGGTTTATGATGATAAGCTGGCTACGGCTCTCTGTTTGACACGTG 557
Db 157 SerPheLeuGluLysPheMetAlaGlyGluValSerTyrGlySerTrpTyrGlnHisVal 176
```

```
Qy 558 CAGAGAGTCTGGGAGACCCGATCGAATCGAATCGTCTTTTCTCAAGTATGAAGACATG 617
Db 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
Qy 618 CATCGGACCTGGTGGATGGTGGACAGCTGCCAGATTCCTCGGGGTCTCTCTGTGAC 677
Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
Qy 678 AAGCCCCAGCTGGAAGCCCTGACGAGCAGCTGC----- 710
Db 217 GluGluThrValAspLeuMetValGluHisThrSerPheLysGluMetLysLysAsnPro 236
Qy 711 -----CACAGCTGTGGACCAAGTCTGCAACGCTGAGGCC 746
Db 237 MetThrAsnTyrThrThrValArgGluPheMetAspHisSerIleSer----- 253
Qy 747 CTGCC---GTGGCGCGGGAAGTGGGTGGGAAGGACATCTTCACCGTCTCCATG 803
Db 254 ---ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGln 272
Qy 804 AATGAGAAGTTTGACTTGGTGTATATAACAGAAAGTGGAAAGTGTGACCTCACGTTT 860
Db 273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 13
US-10-769-507-1
; Sequence 1, Application US/10769507
; Publication No. US20050019788A1
; GENERAL INFORMATION:
; APPLICANT: Squires, James E.
; APPLICANT: Lin, Zhihong E.
; TITLE OF INVENTION: Genetic Markers for Skatole Metabolism
; FILE REFERENCE: P06153US03
; CURRENT APPLICATION NUMBER: US/10/769,507
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: 10/024,628
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: 09/288,037
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/081,037
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 295
; TYPE: PRT
; ORGANISM: porcine
US-10-769-507-1

Alignment Scores:
Pred. No.: 4.22e-29 Length: 295
Score: 456.00 Matches: 98
Percent Similarity: 54.2% Conservative: 51
Best Local Similarity: 35.6% Mismatches: 108
Query Match: 10.4% Indels: 18
DB: 5 Gaps: 3

US-10-768-158-1 (1-2419) x US-10-769-507-1 (1-295)
Qy 90 GCGGTGGCGTCCGCCCTTCTGCGCGGGAAGATGGAGAGATCGCCAACTTCCCGTG 149
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
Qy 150 CGGCCCGACGCTGTGGATCGTACCTACCCAGTCCGCGACCGCTTCTCGCAGGAG 209
Db 37 TrpProaspAspValLeuIleSerThrTyrProLysSerGlyThrTrpValSerGlu 56
Qy 210 GTGGTCTACTTGTGAGCAGCGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
Db 57 IleLeuaspLeuIleTyrGlnGlyAspLeuGlnLysCysGlnArgAlaProIlePhe 76
Qy 270 GAGCAGTCCCGGTCTCTGGAGTACCCACGCGG-----GGCTGGACATCATC 317
```

Dh	77	Val	Arq	Val	Pro	Phe	Leu	Glu	Phe	Leu	Syl	Pro	Arg	Cys	Pro	Thr	Gly	Phe	Glu	Leu	Leu	96
Qy	318	AAG	GA	CT	CC	CA	CT	CC	CG	CG	CT	CAT	CA	AG	GC	CA	CT	GC	CG	CT	TC	377
Dh	97	Lys	Asp	Thr	Pro	Ala	Pro	Arg	Leu	Leu	Syl	Thr	His	Leu	Pro	Leu	Thr	Leu	Leu	Pro	Gln	116
Qy	378	GAC	CT	CC	CA	AA	TG	GAG	AG	CT	CC	AA	GT	CA	TCT	AT	TAT	TG	CG	CA	CC	437
Dh	117	Thr	Leu	Leu	Asp	Gln	Lys	Val	Val	Lys	Val	Tyr	Val	Ala	Ala	Arg	Asn	Ala	Lys	Asp	Val	136
Qy	438	GTG	TCT	TAT	TAT	CAG	TTC	CA	CG	CT	CT	CG	GG	AC	CA	TG	AG	CT	AC	CG	AG	497
Dh	137	Val	Ser	Tyr	Tyr	His	Phe	Tyr	Arg	Met	Ala	Lys	Val	Tyr	Pro	Asn	Pro	Gly	Thr	Trp	Asp	156
Qy	498	GA	AT	TCT	CG	CG	AG	GT	TAT	GA	AT	GA	TG	CG	CT	AC	GG	CT	AC	GG	CT	557
Dh	157	Ser	Phe	Leu	Leu	Glu	Asp	Phe	Met	Ala	Gly	Glu	Val	Ser	Tyr	Gly	Ser	Trp	Tyr	Gln	His	176
Qy	558	CAG	GAG	TTC	TCG	GAG	CA	CG	CA	CT	GG	AC	TG	CT	TTT	TTC	CA	AG	TAT	CA	AA	617
Dh	177	Gln	Glu	Trp	Trp	Glu	Leu	Arg	His	Thr	His	Pro	Val	Leu	Tyr	Leu	Phe	Tyr	Glu	Asp	Met	196
Qy	618	CAT	CGG	CA	CT	CG	TG	TA	CA	GT	GG	TG	AG	CA	GT	CG	CA	AT	TCT	CG	GG	677
Dh	197	Lys	Glu	Asn	Pro	Lys	Arg	Glu	Ile	Gln	Lys	Leu	Glu	Phe	Val	Gly	Arg	Ser	Leu	Pro	216	
Qy	678	AAG	CC	CA	CG	CA	GT	CG	CA	CT	TC	AG	CG	AG	CA	CT	GC	-----	CAC</			

RESULT 14

```

RES001 14
US-10-769-507-3
; Sequence 3, Application US/10769507
; Publication No. US2005001978A1
; GENERAL INFORMATION:
; APPLICANT: Squires, James E.
; APPLICANT: Lin, Zhihong E.
; TITLE OF INVENTION: Genetic Markers for Skatole Metabolism
; FILE REFERENCE: P06153U03
; CURRENT APPLICATION NUMBER: US/10/769,507
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: 10/024,628
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: 09/288,037
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/081,037
; PRIOR FILING DATE: 1998-04-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-769-507-3

```

Alignment Scores:	
Pred. No.:	4, 22e-29
Score:	456.00
Percent Similarity:	54.5%
Best Local Similarity:	33.7%
Mismatches:	101
Conservative:	58
Matches:	94
Length:	295

Query Match:	10.4%	Indels:	26
DB:	5	Gaps:	4
US-10-768-158-1 (1-2419) x US-10-769-507-3 (1-295)			
QY	90	GGCGTGGCGGTCGCCCTTCGGCGGGGAAGATGGAGAGATCCCAACTTCCTCCGGTG	149
Db	17	:::: :::: :::: :::: :::: ::::	36
QY	150	CGGCCACGACGTGTGATCGTCACCTPACCACCAAGTCCGCACCAAGTGTCTGCAGGAG	209
Db	37	:::: :::: :::: :::: :::: ::::	56
QY	210	GTGGTCTACTTTGGTAGCAGCGGCGTCAACCCGATGAGATCGGCTTGATGAACATCGAC	269
Db	57	::::::       ::::::       ::::::	76
QY	270	GAGCAGTCCGGTCTGGAGTACCCACAGCGG-----GGCTGGACATCATC	317
Db	77	:::::            ::::       ::::       ::::	96
QY	318	AAGNACTGACCTCTCCCGCTCATCAAGACCCACCTGCCCTACCGCTTTCGGCCCTCT	377
Db	97	::    ::    ::    ::    ::    ::	116
QY	378	GACCTCCACAATGGAGCTCCAAGTCACTATATATGGCTGCACCAACCCAAAGATCTGGT	437
Db	117	::    ::    ::    ::    ::    ::	136
QY	438	GTGCTATTATATGATGTTCCACCGCTCTCTCGGGACCATGAGCTACCGAGGACCTTTCAA	497
Db	137	::    ::    ::    ::    ::    ::	156
QY	498	GAATTCTCCGAGGTTTATGAATGATAAGCTGGCTACGCTCTGTTGTGAGCACGTG	557
Db	157	::    ::    ::    ::    ::    ::	176
QY	558	CAGAGATTCTGGGAGCACCGCATGGACTCGAACGCTGTTTTTCTCAAGTATGAACACATG	617
Db	177	::    ::    ::    ::    ::    ::	196
QY	618	CATCGGACCTGGTGACGATGGTGAGCAGCTGGCCGATTCCTGGGGTGTCTGTGTGAC	677
Db	197	::    ::    ::    ::    ::    ::	216
QY	678	AAGCCCCAGCTGGAAGCCCTGACGAGCACTGC-----710	
Db	217	::    ::    ::    ::    ::    ::	236
QY	711	-----CACCAGCTGGTGGACAGCTGTGTGCAACCTGAGGCC	746
Db	237	MetThrAsnTyrThrValArgArgGluPheMetAspHisSer-----	253
QY	747	CTGCC---GTGGCGCGGGAAGAGTGGCTGTGGAAGGACATTTACCGTCTCCATG	803
Db	254	::    ::    ::    ::    ::    ::	272
QY	804	AATGAGAAGTTTGATCGTGTATATAAACAAGATCGGAAAGTGTACCTCAGTTT	860
Db	273	::    ::    ::    ::    ::    ::	891
RESULT 15			
US-10-769-507-6			
; Sequence 6, Application US/10769507			
; Publication No. US20050019788A1			
; GENERAL INFORMATION:			
; APPLICANT: Squires, James E.			
; APPLICANT: Lin, Zhihong B.			
; TITLE OF INVENTION: Genetic Markers for Skatole Metabolism			
; FILE REFERENCE: P06153US03			
; CURRENT APPLICATION NUMBER: US/10/769,507			
; CURRENT FILING DATE: 2004-01-30			
; PRIOR APPLICATION NUMBER: 10/024,628			



```
Qy 618 CATCGGACGTGGTACCATGTTGGAGCAGCTGCCAGATTCTCTGGGGTGCTCTGTGAC 677
Db      :::::::::::::::::::: |||
197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
Qy 678 AAGCCCGAGCTGGGAAGCCCTGACGAGCAGCTGC----- 710
Db      :::::::::::::::::::: |||
217 GluGluThrValAspLeuMetValGluHisThrSerPheLysGluMetLysLysAsnPro 236
Qy 711 -----CACCAGCTGGTGGACGAGCTGCTCAACGCTGAGGCC 746
Db      :::::::::::::::::::: |||
237 MetThrAsnTyrThrValArgGluPheMetAspHisSerIleSer----- 253
Qy 747 CTGCCC---GTGGCCGGGGAAGAGTTGGGCTGTGGAGGACATCTTCACCGTCTCCATG 803
Db      :::::::::::::::::::: |||
254 ---ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGln 272
Qy 804 AATGAGAAGTTGACTGGTGTATATAACAGAGAGTGGAAAGTGTGACCTCACGTTT 860
Db      :::::::::::::::::::: |||
273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 17
US-10-287-436A-1246
; Sequence 1246, Application US/10287436A
; Publication No. US20050202421A1
; GENERAL INFORMATION:
; APPLICANT: CHILDREN'S HOSPITAL MEDICAL CENTER
; TITLE OF INVENTION: METHOD FOR DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS
; FILE REFERENCE: 10872.514696
; CURRENT APPLICATION NUMBER: US/10/287,436A
; CURRENT FILING DATE: 2002-10-31
; PRIOR FILING DATE: 2002-10-31
; NUMBER OF SEQ ID NOS: 1446
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1246
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-436A-1246

Alignment Scores:
Pred. No.: 4,22e-29 Length: 295
Score: 456.00 Matches: 94
Percent Similarity: 54.5% Conservative: 58
Best Local Similarity: 33.7% Mismatches: 101
Query Match: 10.4% Indels: 26
DB: 5 Gaps: 4

US-10-768-158-1 (1-2419) x US-10-287-436A-1246 (1-295)
Qy 90 GGCGTGGCTGCGCCCTTCTGCGGGGAAGATGGAGAGATGCCCACTTCCCGGTG 149
Db      :::::::::::::::::::: |||
17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
Qy 150 CGGCCGAGCGAGCTGTGATCGTCACTACCCAGTCCGCGCAGCAGCTTCTCGCAGGAG 209
Db      :::::::::::::::::::: |||
37 ArgProAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpValSerGln 56
Qy 210 GTGGTCTACTTGGTGAGCAGCGGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
Db      :::::::::::::::::::: |||
57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysHisArgAlaProIlePhe 76
Qy 270 GAGCAGCTCCGGTCTGAGTACCCAGCCG-----GGCTGGACATCATC 317
Db      :::::::::::::::::::: |||
77 MetArgValProPheLeuGluPheLysValProGlyIleProSerGlyMetGluThrLeu 96
Qy 318 AAGAACTGACCTTCCCGCTCATCAGAGCCAGCTGCTACCGCTTCTGCTCTCT 377
Db      :::::::::::::::::::: |||
97 LysAsnThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116
Qy 378 GACCTCCACATGGAGACTCCCAAGTCACTATATGCTCGCCCAACCCCAAGGATCTGTG 437
Db      :::::::::::::::::::: |||
```

```
Db 117 ThrLeuLeuAspGlnLysValLysValValTyrValAlaArgAsnAlaLysAspValAla 136
Qy 438 GTCTCTATTATCAGTTCCACCCCTCTCTCGGGACCATGAGCTACCGAGCAGCTTTCAA 497
Db      :::::::::::::::::::: |||
137 ValSerTyrTyrHisPheTyrHisMetAlaLysValTyrProHisProGlyThrTrpGlu 156
Qy 498 GAATTCTCGCGGAGTTTATGAATGATAAGCTGGGCTACCGCTCTCTGGTTTGACGAGTG 557
Db      :::::::::::::::::::: |||
157 SerPheLeuGluLysPheMetAlaGlyValSerTyrGlySerTyrGlnHisVal 176
Qy 558 CAGAGTCTGGGAGCACCACATGGAAGTCTGCTTCTCAAGTATGAAGACATG 617
Db      :::::::::::::::::::: |||
177 GlnGluTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
Qy 618 CATCGGACCTGTGACGATGGTGGAGCAGCTGCCAGATTCCTGGGGTGTCTGTGAC 677
Db      :::::::::::::::::::: |||
197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
Qy 678 AAGCCCGAGCTGGGAAGCCCTGACGAGCAGCTGC----- 710
Db      :::::::::::::::::::: |||
217 GluGluThrValAspLeuMetValGluHisThrSerPheLysGluMetLysLysAsnPro 236
Qy 711 -----CACCAGCTGGTGGACGAGCTGCTCAACGCTGAGGCC 746
Db      :::::::::::::::::::: |||
237 MetThrAsnTyrThrValArgGluPheMetAspHisSerIleSer----- 253
Qy 747 CTGCCC---GTGGCCGGGGAAGAGTTGGGCTGTGGAGGACATCTTCACCGTCTCCATG 803
Db      :::::::::::::::::::: |||
254 ---ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGln 272
Qy 804 AATGAGAAGTTGACTGGTGTATATAACAGAGAGTGGAAAGTGTGACCTCACGTTT 860
Db      :::::::::::::::::::: |||
273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 18
US-10-769-507-2
; Sequence 2, Application US/10769507
; Publication No. US20050019788A1
; GENERAL INFORMATION:
; APPLICANT: Squires, James E.
; TITLE OF INVENTION: Genetic Markers for Skatole Metabolism
; FILE REFERENCE: P06153US03
; CURRENT APPLICATION NUMBER: US/10/769,507
; CURRENT FILING DATE: 2004-01-30
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: 09/288,037
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/081,037
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-769-507-2

Alignment Scores:
Pred. No.: 1,11e-28 Length: 295
Score: 451.00 Matches: 93
Percent Similarity: 54.5% Conservative: 57
Best Local Similarity: 33.8% Mismatches: 107
Query Match: 10.2% Indels: 18
DB: 5 Gaps: 3

US-10-768-158-1 (1-2419) x US-10-769-507-2 (1-295)
Qy 90 GGCGTGGCTGCGCCCTTCTGCGGGGAAGATGGAGAGATGCCCACTTCCCGGTG 149
Db      :::::::::::::::::::: |||
17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
```



```
QY 150 CGGCCACGACGTCGTGTGATCGTACCTACCCCAAGTCCGACACGAGTCTGTCGAGGAG 209
Db 37 ArgProAspAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpValSerGln 56
QY 210 GTGGTCTACTTGGTGAAGCGGCGCTACCCCGATGAGATCGGCTTGATGAACATCGAC 269
Db 57 IleLeuAspMetIleTyrGlnGlyAspLeuGluLysCysHisArgAlaProIlePhe 76
QY 270 GAGCAGCTCCCGGTCTCTGAGTACCCACAGCCG-----GGCCTCGACATCATC 317
Db 77 MetArgValProPheLeuGluPheLysAlaProGlyIleProSerGlyMetGluThrLeu 96
QY 318 AAGGAACGTACTCTCCCGCTCATCAAGAGCCACCTGCGCTACCGTCTTCTGCGCTCT 377
Db 97 LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116
QY 378 GACCTCCACATGAGACTCCCAAGTCACTATATGCTCGCAACCCCAAGAGTCTGGTG 437
Db 117 ThrLeuLeuAspGlnLysValValTyrValAlaAlaArgAsnAlaLysAspValAla 136
QY 438 GTGTCTTATTATCATGTTCCACCGCTCTCTCGCGGACCATGAGCTACCGAGCACCTTTCAA 497
Db 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp 156
QY 498 GAATTTCTCCGAGAGTTTATGAATGATAAGCTGGGCTACGGCTCTGTTTGAGACAGTG 557
Db 157 SerPheLeuGluLysPheMetValGlyGluValSerTyrGlySerTyrGlnHisVal 176
QY 558 CAGGAGTCTTGGGAGCACCGCATCGACTCGAACGTCGTTTCTCAAGTATGAAGACATG 617
Db 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
QY 618 CATCGGACGTGTGACGATGGTGGAGCTGTCGCGGACCATGAGCTACCGAGGACCTTTCAA 677
Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyHisSerLeuPro 216
QY 678 AAGCCCGAGCTGGAAGCCCTGACGAGCATGTC-----CACCAGCTGGTGACACAG 728
Db 217 GluGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236
QY 729 TGCTGCACGCTGAGGCCCTGCC-----GTG 755
Db 237 MetThrAsnTyrThrValProGlnGluPheMetAspHisSerIleSerProPheMet 256
QY 756 GCGCGGGAAGAGTTGGGCTGTGAAGGACATCTTACCGTCTCCATGAATGAGAAGTTT 815
Db 257 ArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGlnAsnGluArgPhe 276
QY 816 GACTTGGTGATAACAGAAAGATGGAAAGTGACCTCAGCTT 860
Db 277 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291
```

## RESULT 19

```
US-10-734-049A-245
; Sequence 245, Application US/10734049A
; Publication No. US20050042624A1
; GENERAL INFORMATION:
; APPLICANT: ITOH, Kyogo
; APPLICANT: SHICHIO, Shigeki
; TITLE OF INVENTION: TUMOR ANTIGEN
; FILE REFERENCE: Q-78382
; CURRENT APPLICATION NUMBER: US/10/734,049A
; PRIOR FILING DATE: 2003-12-12
; PRIOR APPLICATION NUMBER: PCT/JP02/05799
; PRIOR FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: JP 2001/177058
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: JP 2001/250728
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 408
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 245
; LENGTH: 295
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-734-049A-245

Alignment Scores:
Pred. No.: 1,11e-28 Length: 295
Score: 451.00 Matches: 93
Percent Similarity: 54.5% Conservative: 57
Best Local Similarity: 33.8% Mismatches: 107
Query Match: 10.2% Indels: 18
Db: 5 Gaps: 3

US-10-768-158-1 (1-2419) x US-10-734-049A-245 (1-295)
QY 90 GCGCTGGCGGTGTCGCGCTTCTGCGCGGGAAGATGAGGAGATCGCCAACTTCCCGGTG 149
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
QY 150 CGGCCACGACGTCGTGTGATCGTCACTACCCCAAGTCCGCGCACGAGTCTGTCGAGGAG 209
Db 37 ArgProAspAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpValSerGln 56
QY 210 GTGTCTTACTTGGTGACCGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
Db 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysHisArgAlaProIlePhe 76
QY 270 GAGCAGCTCCCGGTCTCTGAGTACCCACAGCCG-----GGCCTCGACATCATC 317
Db 77 MetArgValProPheLeuGluPheLysAlaProGlyIleProSerGlyMetGluThrLeu 96
QY 318 AAGGAACGTACTCTCCCGCTCATCAAGAGCCACCTGCGCTACCGCTTCTGCGCTCT 377
Db 97 LysAspThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116
QY 378 GACCTCCACATGAGACTCCCAAGTCACTATATGCTCGCAACCCCAAGAGTCTGGTG 437
Db 117 ThrLeuLeuAspGlnLysValValTyrValAlaAlaArgAsnAlaLysAspValAla 136
QY 438 GTGTCTTATTATCATGTTCCACCGCTCTCTCGCGGACCATGAGCTACCGAGGACCTTTCAA 497
Db 137 ValSerTyrTyrHisPheTyrHisMetAlaLysValHisProGluProGlyThrTrpAsp 156
QY 498 GAATTTCTCCGAGAGTTTATGAATGATAAGCTGGGCTACGGCTCTGTTTGAGACAGTG 557
Db 157 SerPheLeuGluLysPheMetValGlyGluValSerTyrGlySerTyrGlnHisVal 176
QY 558 CAGGAGTCTTGGGAGCACCGCATCGACTCGAACGTCGTTTCTCAAGTATGAAGACATG 617
Db 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGluAspMet 196
QY 618 CATCGGACGTGTGACGATGGTGGAGCTGTCGCGGACCATGAGCTACCGAGGACCTTTCAA 677
Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyHisSerLeuPro 216
QY 678 AAGCCCGAGCTGGAAGCCCTGACGAGCATGTC-----CACCAGCTGGTGACACAG 728
Db 217 GluGluThrValAspPheMetValGlnHisThrSerPheLysGluMetLysLysAsnPro 236
QY 729 TGCTGCACGCTGAGGCCCTGCC-----GTG 755
Db 237 MetThrAsnTyrThrValProGlnGluPheMetAspHisSerIleSerProPheMet 256
QY 756 GCGCGGGAAGAGTTGGGCTGTGAAGGACATCTTACCGTCTCCATGAATGAGAAGTTT 815
Db 257 ArgLysGlyMetAlaGlyAspTrpLysThrPheThrValAlaGlnAsnGluArgPhe 276
QY 816 GACTTGGTGATAACAGAAAGATGGAAAGTGACCTCAGCTT 860
Db 277 AspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291

RESULT 20
US-10-072-012-586
; Sequence 586, Application US/10072012
```

```

; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkete, Richard
; APPLICANT: Li, Li
; APPLICANT: Ganggoli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lopley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 586
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-586

Alignment Scores:
Pred. No.: 1,65e-28 Length: 302
Score: 449.00 Matches: 96
Percent Similarity: 55.8% Conservative: 52
Best Local Similarity: 36.2% Mismatches: 99
Query Match: 10.2% Indels: 18
DB: 4 Gaps: 3

US-10-768-158-1 (1-2419) x US-10-072-012-586 (1-302)
QY 126 GAGGAGTCCCAACTCTCCGGTCGGCGCAGCGTGTTGATCGTACCTACCCCAAG 185
Db 36 AspylletlrpsPheGlnAlaLysProAspLeuLeuileSerThrTyxPrlys 55
QY 186 TCGGGCACCAGCTTGCTGCAGGAGGTGGTCTACTTGGTGAGCCAGGGGGCTGACCCCAG 245

```





Db 94 LeuProLeuGluLeuLeuProLysSerPheLeuSerSerLysAlaLysIleIleTyrVal 113  
Qy 414 GCTCGCAACCCCAAGGATGCTGTGCTTATTATCAGTTCCACCGCTCTCTCGGACC 473  
Db 114 LeuArgAsnProLysAspValAlaValSerTyrTyrHisPheSerArgSerHisLysAsp 133  
Qy 474 ATG---AGCTACCGAGGACCTTTCAAGAAATTCGCCGAGGTTTATGAATGATAAGCTG 530  
Db 134 LeuProAlaAspProGlyThrPheGluGluPheLeuGluAlaPheLeuAsnGlyLysVal 153  
Qy 531 GGCTACGCTCCTGTTGACACGTCGACGAGTTCGGGAGCACCGCATCGACTCGAAC 590  
Db 154 LeuTyrGlySerTyrPheAspHisValLeuGlyTyrPheGluLeuArgProGluProGln 173  
Qy 591 GTGCTTTTTCATAGTATGAAGACATGCATCGGACCTGTGTCAGCATGTCGGAGCAGCTG 650  
Db 174 ValLeuPheLeuAspTyrGluAspLeuLysGluAspProAlaGlyGluIleLysIle 193  
Qy 651 GCCAGATCTCTGGGGTGTCTCTGTGACAAGCCCGAGCTGGAAGCCCTGACGAGCAGCTGC 710  
Db 194 AlaGluPheLeuGlyLeuProLeuSerGluGluLeuAspLysLeuLeuAspHisSer 213  
Qy 711 CAC-----CAGCTGGTGACAGTGTGTCACGCTGAGGCCCTGCCCGTGGC--- 758  
Db 214 SerPhePheLeuMetLysLeuAsnProLeuSerAsnTyrGluThrLeuCysLeuGlyLys 233  
Qy 759 -----CGGGGAAGAGTTGGCTGTGAAGGACATC 788  
Db 234 SerLysGlyArgLysSerProPheMetArgLysGlyLeuValGlyAspTyrLysAsnTyr 253  
Qy 789 TTCACCTCTCCATGAATGAGAAGTTTCACTTGGTGTATTAACAGAAAG 836  
Db 254 PheThrProGluGlnAsnGluLysPheAspLysValIleLysGluLys 269

## RESULT 24

US-10-072-012-867  
; Sequence 867, Application US/10072012  
; Publication No. US20040033493A1  
; GENERAL INFORMATION:  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Zernhusen, Bryan  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Li, Li  
; APPLICANT: Gangolli, Esha  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Anderson, David W.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Taupier Jr, Raymond J.  
; APPLICANT: Gusev, Vladimir Y.  
; APPLICANT: Colman, Steven D.  
; APPLICANT: Wolenc, Adam R.  
; APPLICANT: Pena, Carol E. A  
; APPLICANT: Furtak, Katarzyna  
; APPLICANT: Grosse, William M.  
; APPLICANT: Alsbrook II, John P.  
; APPLICANT: Lepley, Denise M.  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Burgess, Catherine E.  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-258  
; CURRENT APPLICATION NUMBER: US/10/072,012  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: 60/265,102  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 60/265,514  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,517  
; PRIOR FILING DATE: 2001-01-31

Alignment Scores:  
Pred. No.: 4,17e-28 Length: 269  
Score: 444.00 Matches: 104  
Percent Similarity: 55.9% Conservative: 39  
Best Local Similarity: 40.6% Mismatches: 83  
Query Match: 10.1% Indels: 30  
DB: 4 Gaps: 7

US-10-768-158-1 (1-2419) x US-10-072-012-867 (1-269)

Qy 141 TTCCCGGTGGGCGCCAGCAGCTGTGTGATCGTACCTACCCCAAGTCCGGACAGCTTG 200  
Db 20 PheGlnAlaArgProAspAspValLeuIleAlaGlyTyrProLysSerGlyThrTrp 39  
Qy 201 CTCGAGGAGGTGTCTACTTGGTGAGCCAGGCGCTGAC-----CCCGAT 245  
Db 40 LeuGlnGluIleLeuSerLeuHisProAsnValGlyAspPheGluProSerProSerAsp 59  
Qy 246 GAGATCGGCTTGATGAACATCGACGAGCTCCCGGTCTCTGGAGTAGTACCCACAGCCGGC 305  
Db 60 ProLeuLeuPheArgAsn-----ProTrpLeuGluTyrProLys---Gly 73  
Qy 306 CTGGACATCATCAAGGAAGTCTG-----ACCTCTCCCGCCCTCATCAAGAGCCAC 353  
Db 74 GluAspTyrTyrGluThrLeuLysProMetProSerSerProArgLeuIleLysThrHis 93  
Qy 354 CTGCCCTACCGCTTCTGCGCTCTGACCTCCACATCGAGACTCCCAAGGTCTATATATG 413  
Db 94 LeuProLeuGluLeuLeuProLysSerPheLeuSerSerLysAlaLysIleIleTyrVal 113  
Qy 414 GCTCGCAACCCCAAGGATGCTGTGCTTATTATCAGTTCCACCGCTCTCTCGGACC 473  
Db 114 LeuArgAsnProLysAspValAlaValSerTyrTyrHisPheSerArgSerHisLysAsp 133  
Qy 474 ATG---AGCTACCGAGGACCTTTCAAGAAATTCGCCGAGGTTTATGAATGATAAGCTG 530  
Db 134 LeuProAlaAspProGlyThrPheGluGluPheLeuGluAlaPheLeuAsnGlyLysVal 153  
Qy 531 GGCTACGCTCCTGTTGAGCAGCTGACGAGTTCGGGAGCACCGCATCGACTCGAAC 590  
Db 154 LeuTyrGlySerTyrPheAspHisValLeuGlyTyrPheGluLeuArgProGluProGln 173  
Qy 591 GTGCTTTTTCATAGTATGAAGACATGCATCGGACCTGTGTCAGGAGTGTGAGCAGCTG 650  
Db 174 ValLeuPheLeuAspTyrGluAspLeuLysGluAspProAlaGlyGluIleLysIle 193  
Qy 651 GCCAGATCTCTGGGGTGTCTCTGTGACAAGCCCGAGCTGGAAGCCCTGACGAGCAGCTGC 710

Db	194	AlaGluPheLeuGlyLeuProLeuSerGluGluGluLeuAspLeuLeuAspHisSer	213
Qy	711	CAC-----CAGCTGGTGGACCAAGCTGAGCGCCTGCGCCGTGGC---758	
Db	214	SerPhePheLeuMetLysLeuAsnProLeuSerAsnTyrGluThrLeuCysLeuGlyLys	233
Qy	759	-----CGGGGAAGAGTTGGCTGTGGGAAGGACATC	788
Db	234	SerLysGlyArgLysSerProPheMetArgLysGlyLeuValGlyAspTrpLysAsnTyr	253
Qy	789	TTACCGCTCCATGAATGAGAAGTTTGACTTGGTGTATAAACACAAG	836
Db	254	PheThrProGluGlnAsnGluLysPheAspLysValIleLysGluLys	269
RESULT 25			
US	-09-981-353-101		
;	Sequence 101, Application US/09981353		
;	Patent No. US2002160382A1		
;	GENERAL INFORMATION:		
;	APPLICANT: Lasek, Amy W.		
;	APPLICANT: Jones, David A.		
;	TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER		
;	FILE REFERENCE: PA-0038 US		
;	CURRENT APPLICATION NUMBER: US/09/981,353		
;	CURRENT FILING DATE: 2001-10-11		
;	NUMBER OF SEQ ID NOS: 194		
;	SOFTWARE: PERL Program		
;	SEQ ID NO 101		
;	LENGTH: 296		
;	TYPE: PRT		
;	ORGANISM: Homo sapiens		
;	FEATURE:		
;	NAME/KEY: misc feature		
;	OTHER INFORMATION: Incyte ID No. US20020160382A1 4874364CD1		
US	-09-981-353-101		
Alignment Scores:			
Pred No.:	8,51e-28	Length:	296
Score:	440.50	Matches:	94
Percent Similarity:	53.4%	Conservative:	54
Best Local Similarity:	33.9%	Mismatches:	110
Query Match:	10.0%	Indels:	19
DB:	3	Gaps:	3
US	-10-768-158-1 (1-2419) x US-09-981-353-101 (1-296)		
Qy	87	CATCGGTGCGGTGCGCCCTTGTGCGCGGGAAGATGGAGGAGATCCCAACTTCGCG	146
Db	16	HisGlyTyrProMetThrCysAlaPheAlaSerAsnTrpGluLysIleGluGlnPheHis	35
Qy	147	GTGCGGCCACGCGAGCTGTGGATCGTCACCTACCCCAAGTCGGGACACAGCTTGCTGCAG	206
Db	36	SerArgProAspAspIleValIleAlaThrTyrProLysSerGlyThrThrTrpValSer	55
Qy	207	GAGTGGTCTACTTGGTCAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATC	266
Db	56	GluIleIleAspMetIleLeuAsnAspGlyAspIleGluLysCysLysArgGlyPheIle	75
Qy	267	GACGAGCAGCTCCCGGTCTCTGGAGTACCCACACGCGGCGCTG-----GAC	311
Db	76	ThrGluLysValProMetLeuGluMetThrLeuProGlyLeuArgThrSerGlyIleGlu	95
Qy	312	ATCATCAAGGAAGTACCTCTCCCGCGCTCATCAAGAGCCACCTGCGCTTCTG	371
Db	96	GlnLeuGluLysAsnProSerProArgIleValLysThrHisLeuProThrAspLeuLeu	115
Qy	372	CCCTCTGACCTCCCAATGAGACTCCAAAGTCACTATATGCTCGCAACCCCAAGAT	431
Db	116	ProLysSerPheTrpGluAsnAsnCysLysMetIleTyrLeuAlaArgAsnAlaLysAsp	135
Qy	432	CTGGTGGTCTTATTATCAGTTCCACCGCTCTCTGCGGACCATGAGTACCGAGCAC	491

```

Db      67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86
      300 CCG-----GGCTCGACATCATCAAGAACTGACTCTCCCGGCTCATCAAG 347
      87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleVallys 106
      348 AGCCACCTGCGCTACCGCTTTCGCGCTCTGACCTCCCAATGAGAGATCCCAAGGTATC 407
      107 ThrHisLeuProProLysValLeuProAlaSerPheTrpGluLysAsnCysLysMetIle 126
      408 TATATGGCTCGCAACCCCAAGAGTCTGCTGTCTTATATCATGATCCACCGCTCTCTG 467
      127 TyrLeuCysArgAsnAlaLysAspValAlaValSerTyrTyrPheLeuLeuMetIle 146
      468 CGGACCATGACTACCGAGGACCTTTCAAGAAATTCCTCCCGGAGTTTATGAATGAAG 527
      147 ThrSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGlnGlyGln 166
      528 CTGGGCTACGCTCTCTGTTGAGACATGTCAGGAGTTCTGGAGACCGCATGAGATCG 587
      167 ValProTyrGlySerTrpTyrAspHisValLysAlaTrpTrpGluLysSerLysAsnSer 186
      588 AACGTGCTTTTCTCAAGTATGAACATGTCATCGGAGCTGTCGACGATCGTGGAGCAG 647
      187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgGluValVallys 206
      648 CTGGCCAGATTCCTGGGGGTCTCTGTGACAAAGCCAGCTGGAAGCCCTGACCGAGCAC 707
      207 LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleGlnHis 226
      708 TGCCAC-----CAGCTGGTGGACGAC 728
      227 ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrMetMetProGluGlu 246
      729 TGCTGCAAGCTGAGGCCCTGCCCC---GTGGCGGGGGAAGATTTGGCTCTGGAAGCAG 785
      247 MetMetAsnGlnLysValSerProPheMetArgLysGlyIleIleGlyAspTrpLysAsn 266
      786 ATCTTTACCGCTCTCCATGAATGAGAAGTTTGACTTGGTGTATAACAGAAAGATGGAAAG 845
      267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnGlnMetLysAsp 286
      846 TGTGACCTCAGCTTT 860
      287 CysThrValLysPhe 291

```

## RESULT 27

```

US-10-201-525-13
; Sequence 13, Application US/10201525
; Publication No. US2006009631A1
; GENERAL INFORMATION:
; APPLICANT: Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5827.005
; CURRENT APPLICATION NUMBER: US/10/201.525
; PRIORITY FILING DATE: 2002-07-22
; PRIORITY FILING DATE: 09/785,343
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: PCT/US99/16750
; PRIOR FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-201-525-13

Alignment Scores:
Pred. No.: 6,5e-27 Length: 295
Score: 430.00 Matches: 94
Percent Similarity: 53.0% Conservative: 57

```

```

Best Local Similarity: 33.0% Mismatches: 116
Query Match: 9.8% Indels: 18
DB: 5 Gaps: 3

US-10-768-158-1 (1-2419) x US-10-201-525-13 (1-295)

QY      60 GAGTTCAGAGCAAGTACTTCGAGTTCCATGGCGCTGGCGCTTCTGCGCGGG 119
      7 GluTyrTyrGluValPheGlyGluPheArgGlyValLeuMetAspLysArgPheThrLys 26
      120 AAGATGAGGAGATCGCAACTCCCGTGGCGGCCAGCGACGTCGTGGATCGTCACCTAC 179
      27 TyrTrpGluAspValGluMetPheLeuAlaArgProAspAspLeuValIleAlaThrTyr 46
      180 CCCAAGTCCGACACAGCTTGTCTGACGAGGTGTCTACTTGTGTGAGCCAGGCGCTGAC 239
      47 ProLysSerGlyThrThrTrpIleSerGluValValTyrMetIleTyrLysGluGlyAsp 66
      240 CCCGATGAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGTCTCTGGAGTATCCACAG 299
      67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86
      300 CCG-----GGCTGGACATCATCAAGAACTGACTCTCCCGGCTCATCAAG 347
      87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleVallys 106
      348 AGCCACCTGCGCTACCGCTTTCGCGCTCTGACCTCCACAAATGAGAGATCCCAAGGTATC 407
      107 ThrHisLeuProProLysValLeuProAlaSerPheTrpGluLysAsnCysLysMetIle 126
      408 TATATGGCTCGCAACCCCAAGAGTCTGCTGTCTTATATCATGATTCACCGCTCTCTG 467
      127 TyrLeuCysArgAsnAlaLysAspValAlaValSerTyrTyrPheLeuLeuMetIle 146
      468 CGGACCATGACTACCGAGGACCTTTCAAGAAATTCCTCCCGGAGTTTATGAATGAAG 527
      147 ThrSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGlnGlyGln 166
      528 CTGGGCTACGCTCTCTGTTGAGCAGCTGCGAGGATTTCTGGAGACCGCATGAGATCG 587
      167 ValProTyrGlySerTrpTyrAspHisValLysAlaTrpTrpGluLysSerLysAsnSer 186
      588 AACGTGCTTTTCTCAAGTATGAACATGTCATCGGAGCTGTCGACGATCGTGGAGCAG 647
      187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgGluValVallys 206
      648 CTGGCCAGATTCCTGGGGGTCTCTGTGACAAAGCCAGCTGGAAGCCCTGACCGAGCAC 707
      207 LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleGlnHis 226
      708 TGCCAC-----CAGCTGGTGGACGAC 728
      227 ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrMetMetProGluGlu 246
      729 TGCTGCAAGCTGAGGCCCTGCCCC---GTGGCGGGGGAAGATTTGGCTCTGGAAGCAG 785
      247 MetMetAsnGlnLysValSerProPheMetArgLysGlyIleIleGlyAspTrpLysAsn 266
      786 ATCTTTACCGCTCTCCATGAATGAGAAGTTTGACTTGGTGTATAACAGAAAGATGGAAAG 845
      267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnGlnMetLysAsp 286
      846 TGTGACCTCAGCTTT 860
      287 CysThrValLysPhe 291

RESULT 28
US-10-072-012-585
; Sequence 585, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchevnev, Velizar
; APPLICANT: Spytek, Kimberly

```





## US-10-199-330-5

## Alignment Scores:

Pred. No.: 3.1e-26 Length: 304  
 Score: 422.00 Matches: 97  
 Percent Similarity: 52.0% Conservative: 59  
 Best Local Similarity: 32.3% Mismatches: 124  
 Query Match: 9.6% Indels: 20  
 DB: 4 Gaps: 5

US-10-768-158-1 (1-2419) x US-10-199-330-5 (1-304)

```

QY 21 ATGCGGAGAGCGAGCGGAGACCCCGGAGCGGAGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
QY 78 TTCGAGTTCATCGCGTGGCGGCTGCGGCCCTCTCGCGCGGGAAGATGGAGGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys 40
QY 138 AACTTCCGGTGGCGGCGGAGCGAGCGTGTGGATCGTCACTACCCCAAGTCCGGCAGCAGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuIleLeuAlaThrTyrProLysSerGlyThrThr 60
QY 198 TTGCTGCAGGAGTGGTCTACTTGGTGAGCGAGCGGCGCTGACCCCGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
QY 258 ATGAACATCGACGAGCGAGTCCCGTCTCGGAG-----TACCCA-----CAGCCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
QY 303 GGCTTGACATCATCAAGGAACATCGCTCTCCCGCTCATCAAGAGCCACCTCCCTCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
QY 363 CGCTTTTCGCTCTGACCTCCACCAATGGAGATCCCAAGTCACTATATGGTTCGCAAC 422
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
QY 423 CCCAAGGATCTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGCGACCATCAGCTAC 482
Db 141 ProLysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSerPheMetProAsp 160
QY 483 CGAGGCACCTTTCAGGAATTTTCGCGGAGGTTTATGAATGATGAAGCTGGGCTACGGCTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlyGlySer 180
QY 543 TGGTTTGACGACGTGAGGAGTCTCGGAGCACCAGCATGAGCTCAAGAGTCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200
QY 603 AAGTATCAAGACATGATCGGAGCTGGTGCAGTGTGGAGCAGCTGGCAGATTCCTG 662
Db 201 PheTyrGluAspIleLysAspProLysArgGluIleGluLysIleLeuLysPheLeu 220
QY 663 GGGGTGCTGTGACAGCCCGACCTGGAAGCCCTGACGAGCAGCTGCCACCATCGTGTG 722
Db 221 GluLysAspIleSerGluGluIleLeuAsnLysIleIleTyrHisThrSerPheAspVal 240
QY 723 -----GACCATGCTGCAAGCTGAGCGCTGCGCCCTGCGCGGGC----- 758
Db 241 MetLysGlnAsnProMetThrAsnTyrThrThrLeuProThrSerIleMetAspHisSer 260
QY 759 -----CGGGAAGAGTTGGGCTGTGGAGGACATCTTCACCGTCTCC 800
Db 261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAla 280
QY 801 ATGAATGAGAAGTTTGACTTGGTGTATAAACAAGAGATGGAAAGTGTGACCTCACCGTTT 860
Db 281 GlnAsnGluGluPheAspLysAspTyrGlnLysLysMetAlaGlySerThrLeuThrPhe 300

```

RESULT 30

## US-10-199-334-5

; Sequence 5, Application US/10199334  
 ; Publication No. US20030008354A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WOODAGE, Trevor et al.  
 ; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING  
 ; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN  
 ; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF  
 ; FILE REFERENCE: CLO00669DIV-3  
 ; CURRENT APPLICATION NUMBER: US/10199,334  
 ; CURRENT FILING DATE: 2002-07-22  
 ; PRIOR FILING DATE: 2001-03-27, 609,816  
 ; PRIOR APPLICATION NUMBER: 60/192,408  
 ; PRIOR FILING DATE: 2000-03-27  
 ; PRIOR APPLICATION NUMBER: 60/212,725  
 ; PRIOR FILING DATE: 2000-06-20  
 ; PRIOR APPLICATION NUMBER: 09/609,816  
 ; PRIOR FILING DATE: 2000-07-03  
 ; NUMBER OF SEQ ID NOS: 23  
 ; SOFTWARE: Fast-Seq for Windows Version 4.0  
 ; SEQ ID NO 5  
 ; LENGTH: 304  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-199-334-5

## Alignment Scores:

Pred. No.: 3.1e-26 Length: 304  
 Score: 422.00 Matches: 97  
 Percent Similarity: 52.0% Conservative: 59  
 Best Local Similarity: 32.3% Mismatches: 124  
 Query Match: 9.6% Indels: 20  
 DB: 4 Gaps: 5

US-10-768-158-1 (1-2419) x US-10-199-334-5 (1-304)

```

QY 21 ATGCGGAGAGCGAGCGGAGACCCCGGAGCGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
QY 78 TTCGAGTTCATCGCGTGGCGGCTGCGGCCCTCTCGCGCGGGAAGATGGAGGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys 40
QY 138 AACTTCCGGTGGCGGCGGAGCGAGCGTGTGGATCGTCACTACCCCAAGTCCGGCAGCAGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuIleLeuAlaThrTyrProLysSerGlyThrThr 60
QY 198 TTGCTGCAGGAGTGGTCTACTTGGTGAGCGAGCGGCGCTGACCCCGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
QY 258 ATGAACATCGACGAGCGAGTCCCGTCTCGGAG-----TACCCA-----CAGCCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
QY 303 GGCTTGACATCATCAAGGAACATGACCTCTCCCGCTCATCAAGAGCCACCTCCCTCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
QY 363 CGCTTTTCGCTCTGACCTCCACCAATGGAGATCCCAAGTCACTATATGGTTCGCAAC 422
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
QY 423 CCCAAGGATCTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGCGACCATCAGCTAC 482
Db 141 ProLysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSerPheMetProAsp 160
QY 483 CGAGGCACCTTTCAGGAATTTTCGCGGAGGTTTATGAATGATGAAGCTGGGCTACGGCTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlyGlySer 180

```

```

QY 543 TGGTTGACGACGTCAGGAGTTCTGGGAGCACCGCATGACTCGAACGTCGTTTCTC 602
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 TrpPheAspHisValLysGlyTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200
QY 603 AAGTATGAAGACATGTCATCGGACCTGGTACGATGTCGGAGCAGCTGGCCAGATTCCTG 662
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 PheTyrGluAspIleLysLysAspProLysArgGluLeuGluLysPheLeu 220
QY 663 GGGGTGTCCTGTGCACAGGCCAGCTGGAAGCCCTGACGAGGACACTGCCACGAGCTGGTG 722
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
221 GluLysAspIleSerGluGluLeuLeuAsnLysIleIleTyrHisThrSerPheAspVal 240
QY 723 -----GACCAGTCTGCAACGCTGAGGCCCTGCGCCCTGGTGGC----- 758
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 MetLysGlnAsnProMetThrAsnTyrThrLeuProThrSerIleMetAspHisSer 260
QY 759 -----CGGGAAGAGTTGGGCTGTGGAGGACATCTTCACCGTCTCC 800
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAla 280
QY 801 ATGAATGAGAAGTTGACTTGGTGTATATAACAGAAAGTGGAAAGTGTGACCTCAGCTTT 860
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
281 GlnAsnGluGluPheAspLysAspTyrGlnLysLysMetAlaGlySerThrLeuThrPhe 300

RESULT 31
US-10-199-329-5
; Sequence 5, Application US/10199329
; Publication No. US20030166189A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR FILING DATE: 2002-07-22
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-5

Alignment Scores:
Pred. No.: 3,1e-26 Length: 304
Score: 422.00 Matches: 97
Percent Similarity: 52.8% Conservative: 59
Best Local Similarity: 32.3% Mismatches: 124
Query Match: 9.6% Indels: 20
DB: Gaps: 5

US-10-768-158-1 (1-2419) x US-10-199-329-5 (1-304)
QY 21 ATGCGGAGACGCGGCGAGACCCCGAGACCCCGGGGGAGTTCGAG---AGCAAGTAC 77
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
QY 78 TTCAGTTCATCGGTCGGCTCGCGCCCTTCTCGCGGGAAGATGGAGGAGATCGCC 137
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys 40
QY 138 AACATTCGCGTCGCGCGCGAGCGTGTGGATCGTCACCTACCCCAAGTCGCGGACACG 197
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
41 AsnPheGlnAlaLysProAspAspLeuIleLeuAlaThrTyrProLysSerGlyThrThr 60

```

```

QY 198 TTGCTGCAGGAGTGTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTG 257
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
QY 258 ATGAACATCGACGAGCAGCTCCCGGTCTCTGGAG-----TACCCA-----CAGCCG 302
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
QY 303 GGCTGTGACATCATCAAGAACTGACCTCTCCCGGCTCATCAAGAGCACCCTGCTTAC 362
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
101 AspLeuGluPheValLeuGluMetSerProGlnLeuIleLysThrHisLeuProSer 120
QY 363 CGCTTCTCTGCTCTGACCTCCACAATGGAGCTCCAAAGTCTCATCTATATGGTTCGCAAC 422
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValIleValAlaAsn 140
QY 423 CCCAAGGATCGTGGTGTCTTATTATCAGTTCCACCGCTCTCTCTGGGACCATGAGCTAC 482
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
141 ProLysAspCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAsp 160
QY 483 CGAGGACCTTTCAAGAACTCTCCGAGGAGTTTATGAATCATTAAGCTGGGCTAGGCTCC 542
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlyGlySer 180
QY 543 TGGTTTGGACGACGTCGAGGAGTTCTGGGAGCACCGCATGACTCGAACGTCGTTTCTC 602
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 TrpPheAspHisValLysGlyTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200
QY 603 AAGTATGAAGACATGTCATCGGACCTGGTACGATGTCGGAGCAGCTGGCCAGATTCCTG 662
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 PheTyrGluAspIleLysLysAspProLysArgGluLeuGluLysIleLeuLysPheLeu 220
QY 663 GGGGTGTCCTGTGCACAGGCCAGCTGGAAGCCCTGACGAGGACACTGCCACGAGCTGGTG 722
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
221 GluLysAspIleSerGluGluLeuLeuAsnLysIleIleTyrHisThrSerPheAspVal 240
QY 723 -----GACCAGTCTGCAACGCTGAGGCCCTGCGCCCTGGTGGC----- 758
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 MetLysGlnAsnProMetThrAsnTyrThrLeuProThrSerIleMetAspHisSer 260
QY 759 -----CGGGAAGAGTTGGGCTGTGGAGGACATCTTCACCGTCTCC 800
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAla 280
QY 801 ATGAATGAGAAGTTGACTTGGTGTATATAACAGAAAGTGGAAAGTGTGACCTCAGCTTT 860
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
281 GlnAsnGluGluPheAspLysAspTyrGlnLysLysMetAlaGlySerThrLeuThrPhe 300

RESULT 32
US-11-108-875-5
; Sequence 5, Application US/11108875
; Publication No. US20050196802A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-4
; CURRENT APPLICATION NUMBER: US/11/108,875
; CURRENT FILING DATE: 2005-04-19
; PRIOR FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: 10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304

```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-875-5

Alignment Scores:
Pred. No.: 3,1e-26 Length: 304
Score: 422.00 Matches: 97
Percent Similarity: 52.0% Conservative: 59
Best Local Similarity: 32.3% Mismatches: 124
Query Match: 9.6% Indels: 20
DB: 6 Gaps: 5

US-10-768-158-1 (1-2419) x US-11-108-875-5 (1-304)
Qy 21 ATGCGGAGAGCGAGCGGAGACCCAGACCCCGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy 78 TTCAGATTCCATGCGGTGCGGCTCGGCCCTTCGCCGCGGGAAGATGGAGGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys 40
Qy 138 AACTTCCCGTGGCGGCGGAGCGTGTGGATCGTCACTACCCCAAGTCCGGCCAGCAGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTyProLysSerGlyThr 60
Qy 198 TTGCTGCAGGAGTGTGCTACTTGGTGAGCGCGCTGACCCCGGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLeuArg 80
Qy 258 ATGAACATCGACGAGCTCCCGTCTGGAG-----TACCCA-----CAGCCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysLysGluLysPro 100
Qy 303 GGCTGTGACATCATCAGGAAGTCACTCTCCCGCTCATCAGAGCCAGCCCTCCCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
Qy 363 CGCTTTTCGCTCTGACCTCCACAATGGAGACTCCCAAGGTCACTATATGGCTCGCAAC 422
Db 121 HisLeuLeuProProSerIleTrpLysGluAsnCysLysIleValTyValAlaArgAsn 140
Qy 423 CCCAAGATCTGGGTGTCTTATATCAGTTCCACCGCTCTCTGCGGACCATAGGTAC 482
Db 141 ProLysAspCysLeuValSerTyTyHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACCTTCAAGATTCTGCGGAGGTTTATGAATGATAGCTGGGTACGGCTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyGluLysPheMetSerGlyLysValValGlyLys 180
Qy 543 TGGTTTGAGCAGCTGCAGGAGTTCCTGGGAGCACCGCATGGACTCGAAGCTGCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyLeu 200
Qy 603 AAGTATGAACATGCATCGGACCTGGTGACGATGGTGAGCAGCTGGCCAGATTCCTG 662
Db 201 PheTyGluAspIleLysLysAspProLysArgGluIleGluLysIleLeuLysPheLeu 220
Qy 663 GGGGTGCTCTGTACAGGCCACCTGGAGCCCTGAGGAGCACTGCCACCAGCTGGTG 722
Db 221 GluLysAspLeuSerGluGluIleLeuAsnLysIleTyHisThrSerPheAspVal 240
Qy 723 -----GACCAGTGTGCAACGCTGAGGCCCTGCGCGTGGGC----- 758
Db 241 MetLysGlnAsnProMetThrAsnTyThrThrLeuProThrSerIleMetAspHisSer 260
Qy 759 -----CGGGGAAGATTGGGTGTGGAGGACATCTTCACCGTCTCC 800
Db 261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyThrValAla 280
Qy 801 ATGATGAGAAGTTGACTTGGTGTATAAACAGAGATGGGAAGTGTGACCTCACGTTT 860
Db 281 GlnAsnGluGluPheAspLysAspTyGlnLysLysMetAlaGlySerThrLeuThrPhe 300
```

```
RESULT 33
US-09-898-570-40
; Sequence 40, Application US/09898570
; Patent No. US20020123612A1
; GENERAL INFORMATION:
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
; FILE REFERENCE: 15966-776CIP
; CURRENT APPLICATION NUMBER: US/09/898,570
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/198,293
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/198,645
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/210,809
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/199,476
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,025
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/224,610
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/200,024
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/199,880
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/218,591
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/271,814
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/215,855
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 09/839,446
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-898-570-40

Alignment Scores:
Pred. No.: 4,9e-26 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.6% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: 3 Gaps: 4

US-10-768-158-1 (1-2419) x US-09-898-570-40 (1-283)
Qy 81 GAGTTCCATCGCTGGCGCTGCGCCCTTCTGCGCGGAAGATGGAGGATCGCCAAC 140
Db 1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
Qy 141 TTCCCGGTGGCGGCGGAGCGACGTGTGGATCGTCACTACCCCAAGTCCGGCCAGCTTG 200
Db 21 PheGlnAlaLysProAspAspLeuLeuAlaThrTyAlaLysAlaGlyThrTrp 40
Qy 201 CTCAGAGAGTGTCTACTTGTGTGACCGGCGGTGACCCCGATGATGATGATGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
Qy 261 AACATCGACGAGCAGCTCCCGGTCTGGAGTAC-----CCACAGCGG-----GGC 305
Db 61 AsnThrTyAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
```

```
Qy 306 CTGACATCATCAAGGAAGTGGACCTCCCGCCTCATCAAGAGCCACCTGCGCTACCGC 365
Db 1 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
Qy 366 TTTCTGCGCTCTGACCTCCACAATGGAGACTCCAAAGTCTATATATGGCTCGCAACCC 425
Db 101 MetLeuProSerPheTrpLysGluAsnSerLysIleTyrValAlaAAsnAla 120
Qy 426 AAGGATCTGGTGTCTTATTATCATGTTCCACCGCTCTCTGCGGACCATGAGTACCGA 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
Qy 486 GGACACCTTTCAAGAAATCTGCGGAGGTTTATGAATGATAGCTGGCTACGGCTCTGG 545
Db 141 GlyThrLeuGlyGluTyrIleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 160
Qy 546 TTTGAGCAGCTGCGAGGAGTTCTGGGACCGCAGCTGAGCTGCAACGCTGCTTTTCTCAAG 605
Db 161 TyrAspHisValLysGlyTrpTipAspValLysAspGlnHisArgIleLeuTyrLeuPhe 180
Qy 606 TATGAAGACATGATCGGACCTGGTACGATGGTGGAGCAGCTGGCAGATTCCTGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysIleAlaLysPheLeuGlu 200
Qy 666 GTGCTCTGTGACAAAGCCAGCTGGAAGCCCTGACGAGCACTGCCACCACTGGTG-- 722
Db 201 LysAspLeuSerGluGluValLeuAsnLysIleTyrHisThrSerPheAspValMet 220
Qy 723 -----GACCATGCTGCAACGCTGAGCGCCCTGCCC----- 752
Db 221 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240
Qy 753 -----GTGGCGGGGAAGATGGTGGCTGGGAAGACATCTTCACCGTCTCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
Qy 804 AATGAGAAGTTTGACTGGTGTATTAACAGAGAAGATGGAAAGTGTGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 279

RESULT 34
US-09-839-446-40
; Sequence 40, Application US/0983946
; Publication No. US2003050232A1
; GENERAL INFORMATION:
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: 15966-776
; CURRENT APPLICATION NUMBER: US/09/839,446
; CURRENT FILING DATE: 2001-04-19
; PRIORITY APPLICATION NUMBER: 60/198,293
; PRIORITY FILING DATE: 2000-04-19
; PRIORITY APPLICATION NUMBER: 60/198,645
; PRIORITY FILING DATE: 2000-04-20
; PRIORITY APPLICATION NUMBER: 60/210,809
; PRIORITY FILING DATE: 2000-06-09
; PRIORITY APPLICATION NUMBER: 60/199,476
; PRIORITY FILING DATE: 2000-04-26
; PRIORITY APPLICATION NUMBER: 60/200,025
; PRIORITY FILING DATE: 2000-04-26
; PRIORITY APPLICATION NUMBER: 60/224,610
; PRIORITY FILING DATE: 2000-08-11
; PRIORITY APPLICATION NUMBER: 60/200,024
; PRIORITY FILING DATE: 2000-04-26
; PRIORITY APPLICATION NUMBER: 60/199,880
; PRIORITY FILING DATE: 2000-04-26
; PRIORITY APPLICATION NUMBER: 60/218,591
; PRIORITY FILING DATE: 2000-07-17
; PRIORITY APPLICATION NUMBER: 60/271,814
```

```
; PRIOR FILING DATE: 2001-02-27
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-839-446-40

Alignment Scores:
Pred. No.: 4, 9e-26 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.6% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: Gaps: 4

US-10-768-158-1 (1-2419) x US-09-839-446-40 (1-283)
Qy 81 GAGTTCCATGCGCTGCGGCTGCCGCTTCTGCGGGAAGATGGAGAGATCGCCAAC 140
Db 1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
Qy 141 TTCCCGGTGCGGCGCCAGCGACGTGTGGATCGTCACTACCCCAAGTCCGGCACCAGCTTG 200
Db 21 PheGlnAlaLysProAspAspLeuLeuIleAlaThrTyrAlaLysAlaGlyThrThrTrp 40
Qy 201 CTGAGAGAGTGTCTACTTGTGTGAGCCAGGGGCTGACCCCATGAGATGCGGCTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
Qy 261 AACATCGACGACGAGCTCCCGGCTCTGGAGTAC-----CCACAGCG-----GGC 305
Db 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
Qy 306 CTGACATCATCAAGGAAGTGGACCTCTCCCGGCTCATCAAGAGCCACCTGCGCTACCGC 365
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
Qy 366 TTTCTGCGCTCTGACCTCCACAATGGAGACTCCAAAGTGTATATGATAGCTGGCTCTCG 425
Db 101 MetLeuProSerPheTrpLysGluAsnSerLysIleTyrValAlaAAsnAla 120
Qy 426 AAGGATCTGGTGTCTTATTATCATGTTCCACCGCTCTCTGCGGACCATGAGTACCGA 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
Qy 486 GGACACCTTTCAAGAAATCTGCGGAGGTTTATGAATGATAGCTGGCTACGGCTCTCG 545
Db 141 GlyThrLeuGlyGluTyrIleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 160
Qy 546 TTTGAGCAGCTGCGAGGAGTTCTGGGACCGCAGCTGAGCTGCAACGCTGCTTTTCTCAAG 605
Db 161 TyrAspHisValLysGlyTrpTipAspValLysAspGlnHisArgIleLeuTyrLeuPhe 180
Qy 606 TATGAAGACATGATCGGACCTGGTACGATGGTGGAGCAGCTGGCAGATTCCTGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysIleAlaLysPheLeuGlu 200
Qy 666 GTGCTCTGTGACAAAGCCAGCTGGAAGCCCTGACGAGCACTGCCACCACTGGTG-- 722
Db 201 LysAspLeuSerGluGluValLeuAsnLysIleTyrHisThrSerPheAspValMet 220
Qy 723 -----GACCATGCTGCAACGCTGAGCGCCCTGCCC----- 752
Db 221 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240
Qy 753 -----GTGGCGGGGAAGATGGTGGCTGGGAAGACATCTTCACCGTCTCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
Qy 804 AATGAGAAGTTTGACTGGTGTATTAACAGAGAAGATGGAAAGTGTGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 279
```

```
Db 261 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 279
RESULT 35
; Sequence 8, Application US/10199330
; Publication No. US20020182681A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL00669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-8
Alignment Scores:
Pred. No.: 4.9e-26 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.6% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: 4 Gaps: 4
US-10-768-158-1 (1-2419) x US-10-199-330-8 (1-283)
QY 81 GAGTTCCATGCGTGGCGTCCGCCCTTCTGCGCGGAAGATGGAGGATCGCCAAC 140
Db 1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
QY 141 TTCCCGGTGCGGCCGACGACGTGTGGATCGTCACTACCCCAAGTCCGGCACCAGCTTG 200
Db 21 PheGlnAlaLysProAspAspLeuIleAlaThrTyrAlaLysAlaGlyThrThrTrp 40
QY 201 CTGACGAGGTGGTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
QY 261 AACATCGACGACGACGCTCCGGTCTGTGAGTAC-----CCACAGCCG-----GGC 305
Db 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
QY 306 CTGGACATCATCAAGGAACGTACCTCTCCCGCCTCATCAAGAGCCACTGCCCTACCGC 365
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
QY 366 TTCTGCGCCTCTGACCTCCCAAGTGGAGACTCCAAAGTTCATCTATATGGTTCGCAACCCC 425
Db 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAsnAla 120
QY 426 AAGGATCTGGTGGTCTTATATATCAGTTCACCGCTCTCTGCGGACCATGAGTACCGA 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
QY 486 GCCACCTTTCAGAAATCTCGCGGAGGTTTATGAATGATGAAGTGGGTACGGCTCTCGG 545
Db 141 GlyThrLeuGlyGluTyrIleGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 160
QY 546 TTGAGCAGGTGCAGGAGTCTTGGGAGCAGCCGATGGAAGTCTGCTTTTCTCAAG 605
```

```
Db 161 TyrAspHisValLysGlyTrpTrpAspValLysAspGlnHisArgIleLeuTrpPhe 180
QY 606 TATGAAGACATGTCATCGGACCTGCTGACGATGTGGAGCAGCTGGCCAGATTCTCTGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 200
QY 666 GTGCTCTGTGACAAGGCCAGCTGGAAGCCCTGACGAGCAGCTGCCACCACTGGTG--- 722
Db 201 LysAspIleSerGluGluValLeuAsnLysIleIleTyrHisThrSerPheAspValMet 220
QY 723 -----GACCACTGCTGCAACGCTGAGGCCCTGCC----- 752
Db 221 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240
QY 753 -----GTGGCGCGGGAAGAGTGGGCTGTGAAGAGACATCTTCCACGCTCTCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
QY 804 AATGAGAAGTTTGACTTGGTGTATATAACAGAGATGGGAAAGTGTGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 279
RESULT 36
US-10-199-330-9
; Sequence 9, Application US/10199330
; Publication No. US20020182681A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL00669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-9
Alignment Scores:
Pred. No.: 4.9e-26 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.6% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: 4 Gaps: 4
US-10-768-158-1 (1-2419) x US-10-199-330-9 (1-283)
QY 81 GAGTTCCATGCGTGGCGTCCGCCCTTCTGCGCGGAAGATGGAGGATCGCCAAC 140
Db 1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
QY 141 TTCCCGGTGCGGCCGACGACGACGTGTGGATCGTCACTACCCCAAGTCCGGCACCAGCTTG 200
Db 21 PheGlnAlaLysProAspAspLeuIleAlaThrTyrAlaLysAlaGlyThrThrTrp 40
QY 201 CTGACGAGGTGGTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
```

```
QY 261 AACATCGACGAGCAGCTCCGGTCTCGAGTAC-----CCACAGCCG-----GGC 305
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
QY 306 CTGGACATCATCAAGGAAGTCACTCCCGCCCTCATCAAGAGCCACCTGCCCTACCGC 365
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
QY 366 TTTCTGCCCTCTGACCTCCACAATGGAGACTCCAAAGTCACTATATATGGTCGCAACCC 425
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAsnAla 120
QY 426 AAGGATCTGGTGGTCTTATTATCACTCCACCGCTCTCTCGCGACCATGAGCTACCGA 485
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
QY 486 GGCACTCTTCAAGAAATCTCGCGGAGGTTTATGAATGATCAAGCTCGGCTCCCTGG 545
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
141 GlyThrLeuGlyGluTyrIleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 160
QY 546 TTTGAGCAGCTGCGAGGAGTCTCGGAGCAGCCGATGAGCTCAAGCTGCTTTTCTCAAG 605
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
161 TyrAspHisValLysGlyTrpTrpAspValLysAspGlnHisArgIleLeuTyrLeuPhe 180
QY 606 TATGAGACATGATCGGAGCTCGTGACGATGTGGAGCAGCTGGCCAGATTCTCGGG 665
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
181 TyrGluAspMetLysGluAspProLysArgGluIleLysIleAlaLysPheLeuGlu 200
QY 666 GTGCTCTGTCAAGCCCGACCTGGAAGCCCTGAGGAGCAGCTGCCACAGCTGGTG--- 722
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
201 LysAspIleSerGluGluValLeuAsnLysIleIleTyrHisThrSerPheAspValMet 220
QY 723 -----GACAGCTGTGCAAGCTGAGCCCTGGCC----- 752
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
221 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240
QY 753 -----GTGGCCGGGAGAGTGTGGCTGTGGAGGACATCTTCACCGTCTCCATG 803
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
```

## RESULT 37

```
US-10-199-334-8
; Sequence 8, Application US/10199334
; Publication No. US20030008354A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; PRIORITY FILING DATE: 2002-07-22
; PRIORITY APPLICATION NUMBER: 09/609,816
; PRIORITY FILING DATE: 2001-03-27
; PRIORITY APPLICATION NUMBER: 60/192,408
; PRIORITY FILING DATE: 2000-03-27
; PRIORITY APPLICATION NUMBER: 60/212,725
; PRIORITY FILING DATE: 2000-06-20
; PRIORITY APPLICATION NUMBER: 09/609,816
; PRIORITY FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-8
Alignment Scores:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; PRIORITY FILING DATE: 2002-07-22
; PRIORITY APPLICATION NUMBER: 09/609,816
; PRIORITY FILING DATE: 2001-03-27
; PRIORITY APPLICATION NUMBER: 60/192,408
; PRIORITY FILING DATE: 2000-03-27
; PRIORITY APPLICATION NUMBER: 60/212,725
; PRIORITY FILING DATE: 2000-06-20
; PRIORITY APPLICATION NUMBER: 09/609,816
; PRIORITY FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-8
Alignment Scores:
```

```
Pred. No.: 4.9e-26 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 53.6% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: 4 Gaps: 4

US-10-768-158-1 (1-2419) x US-10-199-334-8 (1-283)
QY 81 GAGTTCATGGCTGGCTGGCTGGCCCTTCTGCGCGGGAAGATGGAGGAGATCGCCCAAC 140
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
QY 141 TTCCCGGTGGCGCCAGCAGCTGGTGGATCGTCACTTACCCCAAGTCCGCGCACCGATTG 200
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
21 PheGlnAlaLysProAspAspLeuLeuIleAlaThrTyrAlaLysAlaGlyThrTrp 40
QY 201 CTGACGAGGTGGTCTACTTGTGAGCCAGGGGCTGACCCCGATGATGATCGGCTTGATG 260
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
QY 261 AACATCGACGAGCAGCTCCCGCTCTCGAGTAC-----CCACAGCCG-----GGC 305
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
QY 306 CTGACATCATCAAGGAAGTCACTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGC 365
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
QY 366 TTTCTGCCCTCTGACCTCCACAATGGAGACTCCAAAGTCACTATATATGGTCGCAACCC 425
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAsnAla 120
QY 426 AAGATCTGTGGTGTCTTATTATCACTCCACCGCTCTCTCGGAGCAGCTACCGA 485
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
QY 486 GGCACTCTTCAAGAAATCTCGCGGAGGTTTATGAATGATCAAGCTCGGCTCCCTGG 545
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
141 GlyThrLeuGlyGluTyrIleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 160
QY 546 TTTGAGCAGCTGCGAGGAGTCTCGGAGCAGCCGATGAGCTCAAGCTGCTTTTCTCAAG 605
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
161 TyrAspHisValLysGlyTrpTrpAspValLysAspGlnHisArgIleLeuTyrLeuPhe 180
QY 606 TATGAGACATGATCGGAGCTCGTGACGATGTGGAGCAGCTGGCCAGATTCTCGGG 665
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
181 TyrGluAspMetLysGluAspProLysArgGluIleLysIleAlaLysPheLeuGlu 200
QY 666 GTGCTCTGTGACAGGCCCGACCTGGAAGCCCTGAGGAGCAGCTGCCACAGCTGGTG--- 722
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
201 LysAspIleSerGluGluValLeuAsnLysIleIleTyrHisThrSerPheAspValMet 220
QY 723 -----GACCAGTGTCTGCAAGCTGAGCCCTGGCC----- 752
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
221 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240
QY 753 -----GTGGCCGGGGAAGAGTGTGGCTGTGGAGGACATCTTCACCGTCTCCATG 803
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
QY 804 AATGAGAAAGTTTGACTTGTGTATAAACAAGAGATGGAAAGTGTGACCTCACGTTT 860
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
261 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 279

RESULT 38
US-10-199-334-9
; Sequence 9, Application US/10199334
; Publication No. US20030008354A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
```

; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF

; FILE REFERENCE: CL000669DIV-3  
; CURRENT APPLICATION NUMBER: US/10/199,334  
; CURRENT FILING DATE: 2002-07-22  
; PRIOR APPLICATION NUMBER: 09/609,816  
; PRIOR FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: 60/192,408  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: 60/212,725  
; PRIOR FILING DATE: 2000-06-20  
; PRIOR APPLICATION NUMBER: 09/609,816  
; PRIOR FILING DATE: 2000-07-03  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 9

; LENGTH: 283

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-199-334-9

Alignment Scores:

Pred. No.: 4.9e-26 Length: 283

Score: 419.50 Matches: 91

Percent Similarity: 53.0% Conservative: 57

Best Local Similarity: 32.6% Mismatches: 112

Query Match: 9.5% Indels: 19

DB: 4 Gaps: 4

US-10-768-158-1 (1-2419) x US-10-199-334-9 (1-283)

QY 81 GAGTTCATGCGTGGCTCGGCTCGCCCTTCTGCGCGGGAAGATGGAGGATCGCCAAC 140

DB 1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20

QY 141 TTCCCGGTGGCGCCAGCAGCTGGATCTCACTACCCCAAGTCGGCGGACAGCTTG 200

DB 21 PheGlnAlaLysProAspAspLeuIleAlaThrTyrAlaLysAlaGlyThrTrp 40

QY 201 CTGCAGGAGTGTCTACTTCTGTGAGCCAGCGGCTGACCCCGATGAGATCGGCTTGATG 260

DB 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60

QY 261 AACATCGAGCAGAGCTCCGGCTCTGGAGTAC-----CCACAGCCG-----GGC 305

DB 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80

QY 306 CTGCACATCATCAAGGAAGTCACTCTCCCGCTCATCAAGCGCCACCTGCCCTACCGC 365

DB 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100

QY 366 TTCTGCGCTCTGACCTCCACAATGGAGACTCCAAAGGTCACTATATGGCTTCGCAACCCC 425

DB 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAsnAla 120

QY 426 AAGATCTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATAGCTACCGA 485

DB 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140

QY 486 GGCACCTTTCAAGAATCTCGCGGAGTTTATGATGATAGCTGGGCTACGGCTCCTGG 545

DB 141 GlyThrLeuGlyGluTyrIleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 160

QY 546 TTGAGCAGCTGCAGGAGTTCTCGGAGCACCAGCATCGACTCGAAGCTGCTTTTCTCAAG 605

DB 161 TyrAspHisValLysGlyTrpTrpAspValLysAspGlnHisArgIleLeuTyrLeuPhe 180

QY 606 TATGAAGACATGCATCGGAGCTGGTACGANGTGGAGCAGCTGGCAGATTCTCTGGG 665

DB 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 200

QY 666 GTGTCCTGTGACAGGCCAGCTGGAAGCCCTGACGGAGCATCTCCACCAGCTGGTG--- 722

DB 201 LysAspIleSerGluGluValLeuAsnLysIleIleTyrHisThrSerPheAspValMet 220

QY 723 -----GACCAGTGTCTCAACGCTGAGGCCCTGCC----- 752

DB 221 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240

QY 753 -----GTGGCGCGGGAAGAGTGGCGCTGGGAAGACATCTTCACCGTCTCCATG 803

DB 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260

QY 804 AATGAGAAAGTTGACTTGGTGATATAAACAGAAAGATGGAAAGTGTGACCTCACGTTT 860

DB 261 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 279

RESULT 39

US-10-199-329-8

; Sequence 8, Application US/10199329

; Publication No. US20030166189A1

; GENERAL INFORMATION:

; APPLICANT: WOODAGE, Trevor et al.

; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING

; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN

; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF

; FILE REFERENCE: CL000669DIV-2

; CURRENT APPLICATION NUMBER: US/10/199,329

; CURRENT FILING DATE: 2002-07-22

; PRIOR APPLICATION NUMBER: 09/609,816

; PRIOR FILING DATE: 2001-03-27

; PRIOR APPLICATION NUMBER: 60/192,408

; PRIOR FILING DATE: 2000-03-27

; PRIOR APPLICATION NUMBER: 60/212,725

; PRIOR FILING DATE: 2000-06-20

; PRIOR APPLICATION NUMBER: 09/609,816

; PRIOR FILING DATE: 2000-07-03

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 283

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-199-329-8

Alignment Scores:

Pred. No.: 4.9e-26 Length: 283

Score: 419.50 Matches: 91

Percent Similarity: 53.0% Conservative: 57

Best Local Similarity: 32.6% Mismatches: 112

Query Match: 9.5% Indels: 19

DB: 4 Gaps: 4

US-10-768-158-1 (1-2419) x US-10-199-329-8 (1-283)

QY 81 GAGTTCATGCGTGGCTCGGCTCGCCCTTCTGCGCGGGAAGATGGAGGATCGCCAAC 140

DB 1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20

QY 141 TTCCCGGTGGCGCCAGCAGCTGTGGATCGTCACTACCCCAAGTCGGCGGACAGCTTG 200

DB 21 PheGlnAlaLysProAspAspLeuIleAlaThrTyrAlaLysAlaGlyThrTrp 40

QY 201 CTGCAGGAGTGTCTACTTCTGTGAGCCAGCGGCTGACCCCGATGAGATCGGCTTGATG 260

DB 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60

QY 261 AACATCGAGCAGAGCTCCCGTCTGGAGTAC-----CCACAGCCG-----GGC 305

DB 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80

QY 306 CTGCACATCATCAAGGAAGTCACTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGC 365

DB 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100

QY 366 TTCTGCGCTCTGACCTCCACAATGGAGACTCAAGGTCACTATATGGCTTCGCAACCCC 425

```
Db 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleLeuTyrValAlaArgAsnAla 120
Qy 426 AAGATCTGGTGGTCTTATTATCAGTTCCACCGCTCTCTGGGACCATGAGTACCGA 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
Qy 486 GGCACCTTCAAGAAATCTCGCGGAGGTTTATGAATGATAGCTGGGTACGGCTCTCGG 545
Db 141 GlyThrLeuGlyGluTyrIleGluInPheLysAlaGlyLysValLeuTrpGlySerTrp 160
Qy 546 TTGAGCAGCTGCAGGAGTTCTGGGACGACCGCATGAGCTCAACGTCCTTTTCTCAAG 605
Db 161 TyrAspHisValLysGlyTyrTrpAspValLysAspGlnHisArgIleLeuTyrLeuPhe 180
Qy 606 TATGAAGACATGCATCGGACCTGGTACGATGTGTGAGCAGCTGGGACGATTCCTGGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysIleAlaLysPheLeuGlu 200
Qy 666 GTGTCTCTGTGACAGGCCAGCTGGAAGCCCTGACGGAGCACTGCCACCGTGGTG-- 722
Db 201 LysAspLeuSerGluGluValLeuAsnLysIleLeuTyrHisThrSerPheAspValMet 220
Qy 723 -----GACCAGTGTGCAACGCTGAGCGCCCTGCCC----- 752
Db 221 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240
Qy 753 -----GTGGCCGGGGAAGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
Qy 804 AATGAGAAAGTTTGACTTGGTGTATAAACAGAGAAGATGGAAAGTGTGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 279

RESULT 40
US-10-199-329-9
; Sequence 9, Application US/10199329
; Publication No. US20030168189A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000569DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-9

Alignment Scores:
Pred. No.: 4,9e-26 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.6% Mismatches: 112
Query Match: 9.5% Indels: 19
Db: 4 Gaps: 4

US-10-768-158-1 (1-2419) x US-10-199-329-9 (1-283)
Qy 81 GAGTTCATGGGTGCGGCTGCGGCCCTTCTCGCGGGGAAGATGGAGGAGATCGCCAAC 140
```

```
Db 1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
Qy 141 TTCCGGTGGCGGCGGACGACGTGTGGATCGTACCTACCCAGTCCGCGACACAGCTTG 200
Db 21 PheGlnAlaLysProAspAspLeuIleAlaThrTyrAlaLysAlaGlyThrTrp 40
Qy 201 CTCAGGAGGTGTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
Qy 261 AACATCCAGCAGCAGCTCCCGGTCTGGAGTAC-----CCACAGCGG-----GGC 305
Db 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
Qy 306 CTGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGC 365
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
Qy 366 TTCTGCGCTCTGACCTCCACATGAGACTCCCAAGTCACTATATGCTGCGCAACCCC 425
Db 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleLeuTyrValAlaArgAsnAla 120
Qy 426 AAGATCTGGTGTGCTTATTATCAGTTCCACCGCTCTCTGGGACCATGAGTACCGA 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
Qy 486 GGACACCTTCAAGAAATCTCGCGGAGGTTTATGAATGATAGCTGGGTACGGCTCTCGG 545
Db 141 GlyThrLeuGlyGluTyrIleGluInPheLysAlaGlyLysValLeuTrpGlySerTrp 160
Qy 546 TTGAGCAGCTGCAGGAGTTCTGGGACGACCGCATGAGCTCAACGTCCTTTTCTCAAG 605
Db 161 TyrAspHisValLysGlyTyrTrpAspValLysAspGlnHisArgIleLeuTyrLeuPhe 180
Qy 606 TATGAAGACATGCATCGGACCTGTGTGACGATGTGGAGCAGCTGGCGACATTCCTCGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 200
Qy 666 GTGTCTCTGTGACAGGCCAGCTGGAAGCCCTGACGGAGCACTGCCACCGTGGTG--- 722
Db 201 LysAspLeuSerGluGluValLeuAsnLysIleLeuTyrHisThrSerPheAspValMet 220
Qy 723 -----GACCAGTGTGCAACGCTGAGCGCCCTGCCC----- 752
Db 221 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240
Qy 753 -----GTGGCCGGGGAAGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
Qy 804 AATGAGAAAGTTTGACTTGGTGTATAAACAGAGAAGATGGAAAGTGTGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 279

RESULT 41
US-11-108-875-8
; Sequence 8, Application US/11108875
; Publication No. US20050196802A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-4
; CURRENT APPLICATION NUMBER: US/11/108,875
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: 10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
```



```
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-875-8

Alignment Scores:
Pred. No.: 4,9e-26 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.6% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: 6 Gaps: 4

US-10-768-158-1 (1-2419) x US-11-108-875-8 (1-283)
Qy 81 GAGTTCCATGGCGTGGCGTCCGCCCTTCTCCCGCGGAAGATGGAGGAGATGCCCAAC 140
Db 1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
Qy 141 TTCGGGTGGCGCCAGCAGCTGTGGATGCTACCTACCCCAAGTCGGCAGCAGCTTG 200
Db 21 PheGlnAlaLysProAspAspLeuIleAlaThrTyrAlaLysAlaGlyThrTrp 40
Qy 201 CTGAGGAGGTGGTCTACTTGTGAGCCAGCGCGTGCACCCGATGAGATCGGCTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgala 60
Qy 261 AACATCCAGCAGCAGCTCCCGTCTGGAGTAC-----CCACAGCCG-----GGC 305
Db 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
Qy 306 CTGACATCATCAAGGAAGTACCTCCCGCTCATCAGAGCAGCAGCTGCCCTACCGC 365
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
Qy 366 TTTCGCCCTCTGACCTCCACATCCAGGTCCATCATATATGCTGCGCAACCC 425
Db 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAsnAla 120
Qy 426 AAGATCTGGTGGTCTTATATCAGTCCCGCTCATCAAGAGCAGCAGCTGCCCTACCGC 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
Qy 486 GGCACCTTTCAAGAAATCTCGCGAGGTATTATGAATGATGAAGCTGGGCTCCGCTCGG 545
Db 141 GlyThrLeuGlyGluTyrIleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 160
Qy 546 TTGAGCAGCTGAGGAGTCTGGGAGCAGCCGATGGATCGAAGCTGCTTTTCTCAAG 605
Db 161 TyrAspHisValLysGlyTrpAspValLysAspGlnHisArgIleLeuTyrLeuPhe 180
Qy 606 TATGAACATCATCGGAGCTGGTGGAGCAGCTGGGAGCAGCTGCCAGATTCCTCGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysIleAlaLysPheLeuGlu 200
Qy 666 GTGCTCTGTACAAGCCAGCTGAAGCCCTGACGGAGCAGCTGCCACAGCTGGTG--- 722
Db 201 LysAspIleSerGluValLeuAsnLysIleIleTyrHisThrSerPheAspValMet 220
Qy 723 -----GACCAGTGTGCAAGCTGAGGCCCTGGCC----- 752
Db 221 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240
Qy 753 -----GTGGCGCGGGAAGAGTGTGGGTGTGGAAGGAGCATCTTCACCGCTCCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
Qy 804 AATGAGAAGTTTGACTTGGTGTATAAACAGAGATGGGAAAGTGTGACCTCACGTTT 860
```

```
Db 261 SerGluAspPheAspGluAspTyrArgLysMetAlaGlySerAsnIleThrPhe 279
US-11-108-875-9
RESULT 42
; Sequence 9, Application US/11108875
; Publication No. US20050196802A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669DIV-4
; CURRENT APPLICATION NUMBER: US/11108,875
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: 10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-875-9
```

```
Alignment Scores:
Pred. No.: 4,9e-26 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.6% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: 6 Gaps: 4

US-10-768-158-1 (1-2419) x US-11-108-875-9 (1-283)
Qy 81 GAGTTCCATGGCGTGGCGTCCGCCCTTCTCCCGCGGAAGATGGAGGAGATGCCCAAC 140
Db 1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
Qy 141 TTCGGGTGGCGCCAGCAGCTGTGGATGCTACCTACCCCAAGTCGGCAGCAGCTTG 200
Db 21 PheGlnAlaLysProAspAspLeuIleAlaThrTyrAlaLysAlaGlyThrTrp 40
Qy 201 CTGAGGAGGTGGTCTACTTGTGAGCCAGCGCGTGCACCCGATGAGATCGGCTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgala 60
Qy 261 AACATCCAGCAGCAGCTCCCGTCTGGAGTAC-----CCACAGCCG-----GGC 305
Db 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
Qy 306 CTGACATCATCAAGGAAGTACCTCCCGCTCATCAAGAGCAGCAGCTGCCCTACCGC 365
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
Qy 366 TTTCGCCCTCTGACCTCCACATCCAGGTCCATCATATATGCTGCGCAACCC 425
Db 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAsnAla 120
Qy 426 AAGATCTGGTGGTCTTATATCAGTCCCGCTCATCAAGAGCAGCAGCTGCCCTACCGC 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
Qy 486 GGCACCTTTCAAGAAATCTCGCGAGGTATTATGAATGATGAAGCTGGGCTCCGCTCGG 545
Db 141 GlyThrLeuGlyGluTyrIleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 160
```

```
Qy 546 TTTCGACGCTGAGGAGTTCTGGGACGACCGCATGGACTCGAAAGTGGCTTTTCTCAAG 605
Db 161 TyrAspHisValysGlyTrpAspVallysAspGlnHisArgIleLeuTyPhe 180
Qy 606 TATGAAGACATGCATCGGACCTGGTGACGATGGGAGCAGCTGGCGAGATTCTCTGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 200
Qy 666 GTGCTCTGTGACAGCCGACCTGGAAGCCCTGACGGAGCAGCTGCCACCACTGGTG--- 722
Db 201 LysAspIleSerGluGluValLeuAsnLysIleTyrHisThrSerPheAspValMet 220
Qy 723 -----GACCACTGCTGCAACGCTGAGCGCCCTGCC----- 752
Db 221 LysGluAsnProMetAlaAsnTyThrThrLeuProSerSerIleMetAspHisSerIle 240
Qy 753 -----GTGGCGCGGGGAAGATTGGGCTGTGGAGGACATCTTCCACGCTCTCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
Qy 804 AATGAGAAGTTTGACTTGGTGTTATAACAGAAAGATGGGAAAGTGTGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 279
RESULT 43
US-10-072-012-583
; Sequence 583, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zethusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
```

```
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 583
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-072-012-583
Alignment Scores:
Pred. No.: 5,03e-26 Length: 304
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.6% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: 4 Gaps: 4
US-10-768-158-1 (1-2419) x US-10-072-012-583 (1-304)
Qy 81 GAGTTCCATGCGGTGCGGCTCCGCGCTTCTGCCGCGGAGATGGAGAGATCGCCAAAC 140
Db 22 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 41
Qy 141 TTCCCGGTGGCGCCGACGAGCTGGATCGTCACCTACCCCAAGTCCGCGACAGCTTG 200
Db 42 PheGlnAlaLysProAspAspLeuLeuIleAlaThrTyrAlaLysAlaGlyThrThrTrp 61
Qy 201 CTGCAGGAGGTGCTACTTGTGGTGGACGAGCGGCTGACCCCGATGAGATCGGCTTCATG 260
Db 62 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 81
Qy 261 AACATCGACGACGACGCTCCCGGCTCCTGGAGTAC-----CCACAGCGC-----GGC 305
Db 82 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 101
Qy 306 CTGGACATCATCAAGAACTGACCTCTCCCGGCTCATCAAGAGCCACCTGCCCTACCGC 365
Db 102 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 121
Qy 366 TTCTGCGCCTCTGACCTCCCAATGGAGACTCCAAAGTCTATATATGGCTCGCAACCCC 425
Db 122 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAsnAla 141
Qy 426 AAGGATCTGGTGGTGCTTATTATTCAGTTCCACCGCTCTCTCGGAGCACCATGAGTACCGA 485
Db 142 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 161
Qy 486 GGCACCTTTCAAGAAATTCGCGGAGGTTTATGATGATAGCTGGGCTACGGCTCCTGG 545
Db 162 GlyThrLeuGlyGluTyrIleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 181
Qy 546 TTTCGACGCTGACGAGGTTCTGGGACGACCGCATGGACTCGAAAGTGGCTTTTCTCAAG 605
Db 182 TyrAspHisValysGlyTrpAspVallysAspGlnHisArgIleLeuTyPhe 201
Qy 606 TATGAAGACATGCATCGGACCTGGTGACGATGGGAGCAGCTGGCGAGATTCTCTGGG 665
Db 202 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 221
Qy 666 GTGCTCTGTGACAGCCGACCTGGAAGCCCTGACGGAGCAGCTGCCACCACTGGTG--- 722
Db 222 LysAspIleSerGluGluValLeuAsnLysIleTyrHisThrSerPheAspValMet 241
Qy 723 -----GACCACTGCTGCAACGCTGAGCGCCCTGCC----- 752
Db 242 LysGluAsnProMetAlaAsnTyThrThrLeuProSerSerIleMetAspHisSerIle 261
Qy 753 -----GTGGCGCGGGGAAGATTGGGCTGTGGAGGACATCTTCCACGCTCTCCATG 803
Db 262 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 281
```



```

Db      41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
Qy      261 AACATCGACGACGAGCTCCCGGTCTCTGGAGTACCCACAGCCG-----GGC 305
Db      61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProProLeuAsnSerGly 80
Qy      306 CTGGACATCATCAAGGAACCTACCTCCCGCTCATCAAGACCCACCTGCCCTACCGC 365
Db      81 LeuAspLeuAlaAenLysMetProSerProArgThrLeuLysThrHisLeuProValGln 100
Qy      366 TTTCTGCGCTCTGACCTCCACAATGGAGACTCCAAAGGTCACTATATGGCTCGCAACCC 425
Db      101 MetLeuProProSerPheTrpLysGluAsnSerGlnIleLeuValAlaArgAsnAla 120
Qy      426 AAGGATCTGGTGTGTTTATATCACTTCACCGCTCTCTCGGACCATGAGTACCGA 485
Db      121 LysAspCysLeuValSerTyrTyrPheSerArgMetAenLysMetLeuProAspPro 140
Qy      486 GGCACCTTTCAGAAATCTGCGGAGGTATTATGAATATGAATAGCTGGGCTACGGCTCTGG 545
Db      141 GlyThrLeuGlyGluTyrIleGluThrPheLysAlaGlyLysValLeuTrpGlySerTrp 160
Qy      546 TTTGAGCAGCTGCAGGAGTTCTGGGAGCCAGCCGATGACTCGAAGCTGCTTTTCTCAAG 605
Db      161 TyrAspHisValLysGlyTrpPheAspValLysAspLysHisArgIleLeuTyrLeuPhe 180
Qy      606 TATGAAGACATGTCATCGGACCTCGTGACGATGGTGGAGCAGCTGCAGATCTCTCGGG 665
Db      181 TyrGluAspMetLysGluAspProLysArgGluIleLysValLysPheLeuGlu 200
Qy      666 GTGTCCTGTGACAGGCCAGCTGATGAAGCCCTGACGAGCACTGCCACCAGCTGGT--- 722
Db      201 LysAspIleSerGluGluValLeuAsnLysIleIleHisThrSerPheAspValMet 220
Qy      723 -----GACCAAGTCTGCAAGCTGAGGCGCTGCCC----- 752
Db      221 LysGlnAsnProMetAlaAsnTyrThrLeuProSerSerIleMetAspHisSerIle 240
Qy      753 -----GTGGCCGGGGAAGATTGGGTGTGGGAAGGACATCTTCACCGTCTCCATG 803
Db      241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
Qy      804 AATCAGAACTTGACTGGTGTTATAACAGAAAGATGGAAAGTGTGACCTCACGTTT 860
Db      261 SerGluAspPheAspGluAspTyrArgLysLysMetAlaGlySerThrIleThrPhe 279

```

## RESULT 46

```

US-10-199-329-10
; Sequence 10, Application US/10199329
; Publication No. US20030166189A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR FILING DATE: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens

```

## US-10-199-329-10

## Alignment Scores:

```

Pred. No.:      4,13e-25      Length:      283
Score:          408.50      Matches:      88
Percent Similarity: 52.0%      Conservat: 57
Best Local Similarity: 31.5%      Mismatches: 115
Query Match:      9.3%      Indels:      19
DB:              4      Gaps:      3

```

## US-10-768-158-1 (1-2419) x US-10-199-329-10 (1-283)

```

Qy      81 GAGTTCATCGCTGCGGTGCGCCCTTCTGCGGGGAAGATGGAGGATCGCCAAC 140
Db      1 GluValAsnGlyIleLeuMetSerLysMetMetSerGluAsnTrpAspLysIleTrpAsn 20
Qy      141 TTCCCGGTGCGGGCCAGCGACGTGTGGATCGTCACTTACCCCAAGTCCGGCACCAGCTTG 200
Db      21 PheGlnAlaLysProAspAspLeuLeuIleAlaThrTyrAlaLysAlaGlyThrThrTrp 40
Qy      201 CTGCAGAGGTGTCTACTTGTGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATG 260
Db      41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
Qy      261 AACATCGACGACGAGCTCCCGGTCTCTGGAGTACCCACAGCCG-----GGC 305
Db      61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProProLeuAsnSerGly 80
Qy      306 CTGGACATCATCAAGGAACCTGACCTCTCCCGCTCATCAAGACCCACCTGCCCTACCGC 365
Db      81 LeuAspLeuAlaAenLysMetProSerProArgThrLeuLysThrHisLeuProValGln 100
Qy      366 TTTCTGCGCTCTGACCTCCACAATGAGACTCCAAAGTCACTATATGGCTCGCAACCC 425
Db      101 MetLeuProProSerPheTrpLysGluAsnSerGlnIleIleTyrValAlaArgAsnAla 120
Qy      426 AAGGATCTGGTGTGTTTATATCACTTCACCGCTCTCTCGGACCATGAGTACCGA 485
Db      121 LysAspCysLeuValSerTyrTyrPheSerArgMetAenLysMetLeuProAspPro 140
Qy      486 GGCACCTTTCAGAAATCTGCGGAGGTATTATGAATATGAATAGCTGGGCTACGGCTCTCG 545
Db      141 GlyThrLeuGlyGluTyrIleGluThrPheLysAlaGlyLysValLeuTrpGlySerTrp 160
Qy      546 TTTGAGCAGCTGCAGGAGTTCTGGGAGCCAGCCGATGACTCGAAGCTGCTTTTCTCAAG 605
Db      161 TyrAspHisValLysGlyTrpTrpAspValLysAspLysHisArgIleLeuTyrLeuPhe 180
Qy      606 TATGAAGACATGTCATCGGACCTGCTGACGATGGTGGAGCAGCTGCGGACATCTCTCGGG 665
Db      181 TyrGluAspMetLysGluAspProLysArgGluIleLysValLysPheLeuGlu 200
Qy      666 GTGTCCTGTGACAGGCCAGCTGGAAGCCCTGACGAGCACTGCCACCAGCTGGT--- 722
Db      201 LysAspIleSerGluGluValLeuAsnLysIleIleHisThrSerPheAspValMet 220
Qy      723 -----GACCAAGTCTGCAAGCTGAGGCGCTGCCC----- 752
Db      221 LysGlnAsnProMetAlaAsnTyrThrLeuProSerSerIleMetAspHisSerIle 240
Qy      753 -----GTGGCCGGGGAAGATTGGGTGTGGGAAGGACATCTTCACCGTCTCCATG 803
Db      241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
Qy      804 AATCAGAACTTGACTGGTGTTATAACAGAAAGATGGAAAGTGTGACCTCACGTTT 860
Db      261 SerGluAspPheAspGluAspTyrArgLysLysMetAlaGlySerThrIleThrPhe 279

```

## RESULT 47

```

US-11-108-875-10
; Sequence 10, Application US/11108875
; Publication No. US20050196802A1
; GENERAL INFORMATION:

```

```
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-4
; CURRENT APPLICATION NUMBER: US/11/108,875
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: 10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-875-10

Alignment Scores:
Pred. No.: 4.13e-25 Length: 283
Score: 408.50 Matches: 88
Percent Similarity: 52.0% Conservative: 57
Best Local Similarity: 31.5% Mismatches: 115
Query Match: 9.3% Indels: 19
DB: 6 Gaps: 3

US-10-768-158-1 (1-2419) x US-11-108-875-10 (1-283)
QY 81 GAGTTCATCGCGTGGCGCTGCGCCCTTCTGCGCGGAAGATGGAGAGATCGCCAAAC 140
DB 1 GluValAsnGlyIleLeuMetSerLysMetMetSerGluAsnTrpAspLysIleTrpAsn 20
QY 141 TTCCCGTGGCGGCCAGCGAGTGTGGTACCTACCTACCCCAAGTCCGGCAGCAGTTG 200
DB 21 PheGlnAlaLysProAspAspLeuLeuIleAlaThrTyralLysAlaGlyThrThrTrp 40
QY 201 CTGCAGAGTGTCTACTTGGTGAGCAGCGGCCCTGACCCCGATGAGTGGGTTGATG 260
DB 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
QY 261 AACATCGACGAGCAGCTCCGGTCTCTGAGTACCCACAGCGC-----GGC 305
DB 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProProLeuAsnSerGly 80
QY 306 CTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGC 365
DB 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValGln 100
QY 366 TTCTGCGCTCTGACCTCCACAAATGGAGATCCAGGTCACTATATGGTTCGCAACCCC 425
DB 101 MetLeuProProSerPheTrpLysGluAsnSerGlnIleIleTyValAlaArgAsnAla 120
QY 426 AAGATCTGGTGGTGTCTATTATCAGTTCCACCGCTCTCTGCGGACCATGAGTACCGA 485
DB 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
QY 486 GGCACTTTCAAGAAATTTGCGCGAGGTTTATGAATGATAGCTGGGCTACGGCTCTCTGG 545
DB 141 GlyThrLeuGlyGlyTyrlleGluThrPheLysAlaGlyLysValLeuTrpGlySerTrp 160
QY 546 TTTGAGCAGCTGCAGGAGTCTTGGGACACCGCATGAGTCAACAGTGTCTTTTCTCAAG 605
DB 161 TyrAspHisValLysGlyTrpTrpAspValLysAspLysHisArgIleLeuTyLeuPhe 180
QY 606 TATCAAGACATGATCGGACCTGGTGACGATGGTGGAGAGCTGGCAGATTCCTCGGG 665
DB 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleValLysPheLeuGlu 200
```

```
QY 666 GTGTCTCTGTGACAAAGGCCCGAGCTGGGAAGCCCTGACGGAGCAGCTGCCACAGCTGGTG--- 722
DB 201 LysAspIleSerGluGluValLeuAsnLysIleHisThrSerPheAspValMet 220
QY 723 -----GACCAGTGTGCAACGCTGAGGCCCTGCCCC----- 752
DB 221 LysGlnAsnProMetAlaAsnTyThrThrLeuProSerSerIleMetAspHisSerIle 240
QY 753 -----GTGGCGCGGGAAGAGTTGGCTGTGAAGAGCATCTTTCACCGCTCTCCATG 803
DB 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyThrThrValAlaGln 260
QY 804 AATGAGAAGTTTCACCTTGGTGTATTAACAGAGATCGGAAGTGTGACCTCAGCTTT 860
DB 261 SerGluAspPheAspGluAspTyArgLysLysMetAlaGlySerThrIleThrPhe 279

RESULT 48
US-10-072-012-584
; Sequence 584, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Mesra
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Beha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 584
; LENGTH: 304
```

```
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-072-012-584

Alignment Scores:
Pred. No.: 4,24e-25 Length: 304
Score: 408.50 Matches: 88
Percent Similarity: 52.0% Conservative: 57
Best Local Similarity: 31.5% Mismatches: 115
Query Match: 9.3% Indels: 19
DB: 4 Gaps: 3

US-10-768-158-1 (1-2419) x US-10-072-012-584 (1-304)
Qy 81 GAGTTCATGCGTGGCGGCTGCCCGCCCTTCGCGCGGGAAGATGGAGGAGATGCCCAAC 140
Db 22 GluValAsnGlyIleLeuMetSerLysMetMetSerGluAsnTrpAspLysIleTrpAsn 41
Qy 141 TTCCCGGTGGCGCCAGCGAGCTGTGGATGTCACCTACCCCAAGTCCCGGACCCAGCTTG 200
Db 42 PheGlnAlaLysProAspAspLeuIleAlaThrTyrAlaLysAlaGlyThrTrp 61
Qy 201 CTGACGAGGTGTCTACTTGTGTAGCCAGGCGCTGACCCCGATGATCGGCTTGATG 260
Db 62 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 81
Qy 261 AACATCAGCAGCAGCTCCCGGTCTGTGAGTACCACAGCGG-----GGC 305
Db 82 AsnThrTyrAspArgHisProPheIleGluThrTrpLeuProProLeuAsnSerGly 101
Qy 306 CTGACATCATCAAGGAACCTGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGC 365
Db 102 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValGln 121
Qy 366 TTCTGCGCCTCTGACCTCCACAATGGAGACTCCAAAGTTCATCTATGGTTCGCAACCCC 425
Db 122 MetLeuProProSerPheTrpLysGluAsnSerGlnIleIleTyrValAlaArgAsnAla 141
Qy 426 AAGATCTGGTGGTCTTATATCAGTTCACCGCTCTCTGCGGACCATGAGTACCGA 485
Db 142 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 161
Qy 486 GGCACCTTCAAGAAATCTGCGGAGGTTTATGATGATAGCTGGGTCCCGTCTCTGG 545
Db 162 GlyThrLeuGlyGluTyrIleGluThrPheLysAlaGlyLysValLeuTrpGlySerTrp 181
Qy 546 TTGAGCAGCTGCAGGAGTCTCGGAGCAGCCGATGCACTGCAACGTCCTTTTCTCAAG 605
Db 182 TyrAspHisValLysGlyTyrTrpAspValLysAspLysHisArgIleLeuTyrLeuPhe 201
Qy 606 TATGAACATGATCGGACCTGGTACGATGTGGAGCAGCTGGCCAGATTCCTGGGG 665
Db 202 TyrGluAspMetLysGluAspProLysArgGluIleLysIleValLysPheLeuGlu 221
Qy 666 GTGTCCTGTGCAAGGCCACCTGGAAGCCCTGAGGAGCACTGCCACGCTGGT--- 722
Db 222 LysAspLysSerGluGluValLeuAsnLysIleIleHisHisThrSerPheAspValMet 241
Qy 723 -----GACCATGTCTGCAAGCTGAGGCCCTGCGCC----- 752
Db 242 LysGlnAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 261
Qy 753 -----GTGGCGCGGGAAGATTGGGTCTGGAAGGACATCTTCACCGTCTCCATG 803
Db 262 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 281
Qy 804 AATGAGAAGTTTGAATCTGTATAAACAGAAAGATGGAAAGTGTGACCTCACCCTTT 860
Db 282 SerGluAspPheAspGluAspTyrArgLysLysMetAlaGlySerThrIleThrPhe 300

RESULT 49
US-10-199-330-6
; Sequence 6, Application US/10199330
```

```
; Publication No. US20020182681A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-6

Alignment Scores:
Pred. No.: 5,67e-25 Length: 304
Score: 407.00 Matches: 94
Percent Similarity: 51.7% Conservative: 61
Best Local Similarity: 31.3% Mismatches: 125
Query Match: 9.2% Indels: 20
DB: 4 Gaps: 5

US-10-768-158-1 (1-2419) x US-10-199-330-6 (1-304)
Qy 21 ATGCGGAGAGCAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy 78 TTGAGTTCATCGGTGGCGGCTGCGCGCTTCTGCGCGGGAAGATGGAGGAGATGCC 137
Db 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys 40
Qy 138 AACTTCCCGTGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 197
Db 41 AsnPheGlnAlaLysProAspAspLeuIleAlaThrTyrProLysSerGlyThrThr 60
Qy 198 TTGCTGCGAGGAGTGTCTACTTGGTGAAGCCAGGCGCTGACCCCGATGAGATCGGCTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
Qy 258 ATGAACATCAGCAGCAGCTCCCGGCTCGAG-----TACCCA-----CAGCCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysGluLysPro 100
Qy 303 GGCTGACATCATCAAGGAAGTACCTCTCCCGCTCATCAAGAGCCACCTGCCCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
Qy 363 CGCTTTTCTGCGCTCTGACCTCCCAATGGAGATCCCAAGGTCTATATATGGTTCGCAAC 422
Db 121 HisLeuIleProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
Qy 423 CCCAGGATCTGGTGGTGTCTTATTATCAGTTCCACCGCTCTCTGCGGACCATGAGTAC 482
Db 141 ProLysAspCysLeuValSerTyrTyrHisPheHisMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACCTTCAAGAAATCTGCGGAGGTCTTATGAATGATGAAGCTGGGCTACGGTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyrGlyLysPheMetSerGlyLysValGlyGlySer 180
Qy 543 TGGTTTGAGCAGCTGCAGGAGTCTTGGAGCAGCCGATGAGCTGCAACGCTGCTTTCTC 602
Db 543 TGGTTTGAGCAGCTGCAGGAGTCTTGGAGCAGCCGATGAGCTGCAACGCTGCTTTCTC 602
```

```
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspThrHisArgIleLeuTyrLeu 200
Qy 603 AAGTATGAAGACATGCGGACCTGCTGACGATGGTGGAGCAGCTGGCCAGATTCCTG 662
Db 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy 663 GGGGTGCTCTGTGACAGCCAGCTGGAAGCCCTGACGGAGCACTGCCACAGCTGGTG 722
Db 221 GluLysThrLeuSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
Qy 723 -----GACAGTGTGCAAGCTGAGCGCTGAGCCCTGCCCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
Qy 753 -----GTGGCGGGGAGAGTTGGGCTGTGGAAGCAGCATCTTCACCGTCTCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
Qy 801 ATGAATGAGAAGTTGACTTGGTGATAAACAGAGATGGGAAGTGTGACCTCACGTTT 860
Db 281 MetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300
RESULT 50
US-10-199-334-6
; Sequence 6, Application US/10199334
; Publication No. US20030008354A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT FILING DATE: 2002-07-22
; PRIOR FILING DATE: 2002-07-22
; PRIOR FILING DATE: 2001-03-27
; PRIOR FILING DATE: 2001-03-27
; PRIOR FILING DATE: 2000-03-27
; PRIOR FILING DATE: 2000-06-20
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-6
Alignment Scores:
Pred. No.: 5,67e-25 Length: 304
Score: 407.00 Matches: 94
Percent Similarity: 51.7% Conservative: 61
Best Local Similarity: 31.3% Mismatches: 125
Query Match: 9.2% Indels: 20
DB: 4 Gaps: 5
US-10-768-158-1 (1-2419) x US-10-199-334-6 (1-304)
Qy 21 ATGCGGAGACGAGCGGACAGACCCCGGCGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy 78 TTCGAGTTCATGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 137
Db 21 MetGluValAspGlyValProThrLeuLeuSerLysGluTyrTrpGluLysValCys 40
Qy 138 AACTTCCCGGTGGCGGACGAGCGATGTGGATCGTACCTACCCCAAGTCCGCGACACGAGC 197
Db 41 AsnPheGlnAlaLysProAspLysPheLeuAlaThrTyrProLysSerGlyThrThr 60
Qy 198 TTGTCGACGAGGTGGTCTACTTGGTGTGACCGGCGCTGACCCCGATGAGATCGGCTTG 257
```

```
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
Qy 258 ATGAACATCGACGAGCAGCTCCCGGTCTCTGAG-----TACCCA-----CAGCCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
Qy 303 GGCTGTGACATCATCAAGGAACATGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
Qy 363 CGCTTTCTGCCCTCTGACCTCCCAATGGAGACTCCCAAGGTCACTCATATATGGCTCGCAAC 422
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
Qy 423 CCCAAGGATCTGGTGGTGTCTTATTATCAGTTTCCACCGCTCTCTCGCGACCATGAGTAC 482
Db 141 ProlsAspCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACCTTTCAAGAAATTCGCCGAGGTTTATGAATGATAAGCTGGGTACGGCTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlyGlySer 180
Qy 543 TGGTTTGACGACGTCGAGGAGTTCCTGGAGCACCAGCATGGAAGTCTGAGCTGCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspThrHisArgIleLeuTyrLeu 200
Qy 603 AAGTATGAAGACATCGGACCTCGTACGATGGTGGAGCAGCTGGCCAGATTCCTG 662
Db 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy 663 GGGGTGCTCTGTGACAGGCCAGCTGGAAGCCCTGACGGAGCACTGCCACAGCTGGTG 722
Db 221 GluLysThrLeuSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
Qy 723 -----GACCAGTGTCTGCAAGCTGAGCGCTGAGCCCTGCCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
Qy 753 -----GTGGCGCGGGAAGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
Qy 801 ATGAATGAGAAGTTGACTTGGTGATATAACAGAGATGGGAAGTGTGACCTCACGTTT 860
Db 281 MetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300
RESULT 51
US-10-199-329-6
; Sequence 6, Application US/10199329
; Publication No. US20030166189A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR FILING DATE: 2002-07-22
; PRIOR FILING DATE: 2001-03-27
; PRIOR FILING DATE: 2001-03-27
; PRIOR FILING DATE: 2000-03-27
; PRIOR FILING DATE: 2000-06-20
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
```







```
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-4
Alignment Scores:
Pred. No.: 1,81e-24 Length: 304
Score: 401.00 Matches: 93
Percent Similarity: 51.7% Conservative: 62
Best Local Similarity: 31.0% Mismatches: 125
Query Match: 9.1% Indels: 20
DB: 4 Gaps: 5

US-10-768-158-1 (1-2419) x US-10-199-330-4 (1-304)
QY 21 ATGGCGGAGAGCGAGCGAGACCCAGACCCCGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
QY 78 TTCGAGTTCATCGCGTGGCGTGGCGGCTTCTGCGCGGGAAGATGGAGGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuLeuLeuSerLysGluTrpTrpGluLysValCys 40
QY 138 AACTTCCCGTGGCGGCGGAGCGAGCGTGGATCGTACCTACCCCAAGTCCGGCACCGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTyrProLysSerGlyThrThr 60
QY 198 TTGCTGAGGAGGTGCTTACTTGGTGAGCAGCGGCGTGCACCCGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
QY 258 ATGAACATCGAGCAGCTCCCGTCTCTGAG---TACCCA-----CAGCCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysGluLysPro 100
QY 303 GGCTTGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACTGCCCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
QY 363 CGCTTTCGCGCTCTGACCTCCACATCGAGACTCCAGGTCATCTATATGCTCGCAAC 422
Db 121 HisLeuIleProSerIleTrpLysGluAsnCysLysIleValIleValAlaArgAsn 140
QY 423 CCCAAGGATCTGGTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGCTAC 482

US-10-768-158-1 (1-2419) x US-10-199-334-4 (1-304)
QY 21 ATGGCGGAGAGCGAGCGAGACCCAGACCCCGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
QY 78 TTCGAGTTCATCGCGTGGCGGCTGCGGCTTCTGCGCGGGAAGATGGAGGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuLeuLeuSerLysGluTrpTrpGluLysValCys 40
QY 138 AACTTCCCGTGGCGGCGGAGCGAGCGTGGATCGTACCTACCCCAAGTCCGGCACCGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTyrProLysSerGlyThrThr 60
QY 198 TTGCTGAGGAGGTGCTTACTTGGTGAGCAGCGGCGTGCACCCGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
QY 258 ATGAACATCGAGCAGCTCCCGTCTCTGAG---TACCCA-----CAGCCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysGluLysPro 100
QY 303 GGCTTGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACTGCCCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
QY 363 CGCTTTCGCGCTCTGACCTCCACATCGAGACTCCAGGTCATCTATATGCTCGCAAC 422
Db 121 HisLeuIleProSerIleTrpLysGluAsnCysLysIleValIleValAlaArgAsn 140
QY 423 CCCAAGGATCTGGTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGCTAC 482
```

```

Db      141 ProlysAspCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAsp 160
Qy      483 CGAGGCACTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCC 542
Db      161 ProGlnAenLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlySer 180
Qy      543 TGGTTTGAGCAGCTGCAGGAGTCTCGGAGCACCGCATGGACTCAAGCTGCTTTTCTC 602
Db      181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200
Qy      603 AAGTATGAAGACATGCATCGGACCTGCTGACGATGGTGGAGCAGCTGGCCAGATTCTCG 662
Db      201 PheTyrGluAspIleLysLysAenProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy      663 GGGGTGCTCTGTGCAAGGCCAGCTGGAAGCCCTGACGGAGCACCTGCCACAGCTGGTG 722
Db      221 GlulysThrTrpSerGlyAspValIleAenLysIleValHisHisThrSerPheAspVal 240
Qy      723 -----GACCAGTGTGCAACGCTGAGCCCTGCGCC----- 752
Db      241 MetLysAspAenProMetAlaAenHisThrAlaValProAlaHisIlePheAenHisSer 260
Qy      753 -----GTGGCCGGGAGAGATTGGGCTGTGGAAGGACATCTTCACCGTCTCC 800
Db      261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAenHisPheThrValala 280
Qy      801 ATGAATGAGAAGTTTGACTTGGTGTATATAACAGAAAGATGGAAAGTGTGACCTCACGTTT 860
Db      281 LeuAenGluAenPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAenPhe 300

```

## RESULT 56

```

US-10-199-329-4
; Sequence 4, Application US/10199329
; Publication No. US20030166189A1
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-4

```

```

Alignment Scores:
Pred. No.: 1,81e-24 Length: 304
Score: 401.00 Matches: 93
Percent Similarity: 51.7% Conservative: 62
Best Local Similarity: 31.0% Mismatches: 125
Query Match: 9.1% Indels: 20
DB: 4 Gaps: 5

```

US-10-768-158-1 (1-2419) x US-10-199-329-4 (1-304)

```

Qy      21 ATGGCGGAGAGCGAGCGCGAGACCCCGGAGGAGTTTCGAG---AGCAAGTAC 77
Db      1 MetAlaLysIleGluLysAenAlaProThrMetGluLysLysProGluLeuPheAenIle 20

```

```

Qy      78 TTCAGTTCATGCGCTGCGGCTGCCGCCCTTCTGCGCGGGAAGATGGAGGAGATCGCC 137
Db      21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys 40
Qy      138 AACTTCCCGGTGGCGGCCAGCGCTGGATCGTACCTACCTACCCCAAGTCCGGCCAGCAGC 197
Db      41 AsnPheGlnAlaLysProAspAspLeuIleLeuAlaThrTyrProLysSerGlyThrThr 60
Qy      198 TTGCTCGAGGAGTGTCTACTTGTGAGCGAGCGGCGCTGACCCCGATGAGATCGGTTG 257
Db      61 TrpMetHisGluIleLeuAspMetIleLeuAenAspGlyAspValGluLysCysLysArg 80
Qy      258 ATGAAATCAGCAGCAGCTCCCGCTCTCGAG-----TACCCA-----CAGCCG 302
Db      81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysLysGluLysPro 100
Qy      303 GGCCTGCACATCATCAAGGAACAGCTCTCCCGCTCATCAAGAGCCACCTCCCTCTAC 362
Db      101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
Qy      363 CGCTTTCGCTCTGACCTCCCAATGGAGACTCCAAAGGTCTATATATGGTTCGCAAC 422
Db      121 HisLeuIleProProSerIleTrpLysGluAenCysLysIleValTyrValAlaArgAsn 140
Qy      423 CCAGAGATCTGGTGTGTCTTATTATCAGTTCCACCGCTCTCTCGCGACCATGAGTAC 482
Db      141 ProlysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSerPheMetProAsp 160
Qy      483 CGAGGCACCTTCAAGAAATTCGCGGAGGTTTATGAATGATAGCTGGGCTACGGCTCC 542
Db      161 ProGlnAenLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlyGlySer 180
Qy      543 TGGTTTGAGCAGCTGCAGGAGTTCCTGGAGCACCGCATGGACTCGAAGCTGCTTTTCTC 602
Db      181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200
Qy      603 AAGTATGAAGACATGCATCGGACCTCGGACCTGTCAGCATGGTGGAGCAGCTGGCCAGATTCTCTG 662
Db      201 PheTyrGluAspIleLysLysAenProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy      663 GGGGTGCTGTGCAAGGCCAGCTGGAAGCCCTGACGGAGCACCTGCCACAGCTGGTG 722
Db      221 GlulysThrTrpSerGlyAspValIleAenLysIleValHisHisThrSerPheAspVal 240
Qy      723 -----GACCAGTGTCTGCAACGCTGAGCCCTGCGCC----- 752
Db      241 MetLysAspAenProMetAlaAenHisThrAlaValProAlaHisIlePheAenHisSer 260
Qy      753 -----GTGGCCGGGAGAGATTGGGCTGTGGAAGGACATCTTCACCGTCTCC 800
Db      261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAenHisPheThrValala 280
Qy      801 ATGAATGAGAAGTTTGACTTGGTGTATATAACAGAAAGATGGAAAGTGTGACCTCACGTTT 860
Db      281 LeuAenGluAenPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAenPhe 300

```

## RESULT 57

```

US-10-072-012-582
; Sequence 582, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie

```





```
QY 483 CGAGGCACCTTCAAGATTCTGCGGAGGTTTATGATGATAGCTGGGCTACGGCTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlyArgSer 180
QY 543 TGGTTTGAGCAGCTGCGAGGAGTCTCGGAGCACCGCATGGACTCGAACGTCGCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspThrHisArgIleLeuTyrLeu 200
QY 603 AAGTATGAAGACATGTCATCGGACCTGCTGACGATGCTGGAGCAGCTGGCGAGATTCTCTG 662
Db 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
QY 663 GGGGTGTCCTGTGCACAGGCCACCTGGAAGCCCTGACGAGCACTGCCACCAGCTGGTG 722
Db 221 GluLysThrLeuSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
QY 723 -----GACCAGTGTGCAAGCTGAGGCCCTGGCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
QY 753 -----GTGGCCCGGGAGAGTTGGGCTGTGAGGAGCATCTTACCGTCTCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
QY 801 ATGAATCAGAGTTTGACTGCTGTATTAACAGAGATGGAAAGTGTGACCTCACGTTT 860
Db 281 MetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300
```

## RESULT 60

```
US-11-079-743-9
; Sequence 9, Application US/11079743
; Publication No. US2005018415A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YANG, Junming
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: BURFORD, Neil
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: REDDY, Roopa
; APPLICANT: RING, Huijun Z.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: YAO, Monique G.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: NGUYEN, Danniel B.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: TANG, Y. Tom
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: BANDMAN, Olga
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0007 PCT
; CURRENT APPLICATION NUMBER: US/11/079,743
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US/10/181,108
; PRIOR FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/10/181,108
; PRIOR FILING DATE: 2002-07-11
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 638419CD1
US-11-079-743-9
```

Alignment Scores:

```
Pred. No.: 2,67e-24 Length: 304
Score: 399.00 Matches: 93
Percent Similarity: 51.3% Conservatives: 61
Best Local Similarity: 31.0% Mismatches: 126
Query Match: 9.1% Indels: 20
DB: 6 Gaps: 5

US-10-768-158-1 (1-2419) x US-11-079-743-9 (1-304)

QY 21 ATGCGGAGAGCGAGCGGAGACCCCGACCCCGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
QY 78 TTCGAGTTCATGCGTGGCGCTGCCGCCCTTCTGCCCGGGAAGATGGAGGAGATGCC 137
Db 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys 40
QY 138 AACTTCCCGGTGGCGGCCAGCGACGTGTGATCGTCACCTACCCCAAGTCCGGCACCAGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuIleLeuAlaThrTyrProLysSerGlyThrThr 60
QY 198 TTCTCTGAGGAGTGTGCTACTTTGGTGACCGGCGCTGACCCCGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
QY 258 ATGAACATCGACGAGCAGCTCCCGGTCTCTGGAG-----TACCCA----- 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysGluLysPro 100
QY 303 GGCCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
QY 363 CGTTTCTGCTCTGACCTCCACATATGGAGATCCCAAGTCTATATATGGTTCGCAAC 422
Db 121 HisLeuIleProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
QY 423 CCCAAGGATCTGGTGTCTTATTATCAGTCTCCCGCTCTCTGGGACCATTGAGCTAC 482
Db 141 ProLysAspCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAsp 160
QY 483 CGAGGCACCTTTCAGAAATTCGCGGAGGTTTATGAATGATTAAGCTGGGCTACGGCTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlyArgSer 180
QY 543 TGGTTGAGCAGCTGCGAGGAGTCTGGGAGCACCGCATGGACTCGCAACGTCGTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspThrHisArgIleLeuTyrLeu 200
QY 603 AAGTATGAAGACATGTCATCGGACCTGGTGACCGATGGTGGAGCAGCTGGCAGATTCTCTG 662
Db 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisValLeuGluPheLeu 220
QY 663 GGGGTGTCCTGTGCACAGGCCACGCTGGAAGCCCTGACGAGCACTGCCACCAGCTGGTG 722
Db 221 GluLysThrLeuSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
QY 723 -----GACCAGTGTGCAAGCTGAGGCCCTGGCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
QY 753 -----GTGGCCCGGGAGAGTTGGGCTGTGAGGAGCATCTTACCGTCTCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
QY 801 ATGAATCAGAGTTTGACTGCTGTATTAACAGAGATGGAAAGTGTGACCTCACGTTT 860
Db 281 MetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300

RESULT 61
US-10-072-012-206
; Sequence 206, Application US/10072012
; Publication No. US20040033493A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meeta
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072.012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 206
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-072-012-206

Alignment Scores:
Pred. No.: 3,53e-24 Length: 295
Score: 397.50 Matches: 85
Percent Similarity: 54.2% Conservative: 58
Best Local Similarity: 32.2% Mismatches: 102
Query Match: 9.0% Indels: 19
DB: 4 Gaps: 4

US-10-768-158-1 (1-2419) x US-10-072-012-206 (1-295)
Qy 126 GAGGAGTCGCAACTCCGGTGGCGGCGGCGAGCGTGTGGATCGTCACCTACCCCAAG 185
Dy 28 GluysValCysAsnPheGlnAlaLysProAspLeuIleLeuAlaThrTyrProLys 47
Qy 186 TCCGGCACCACTGTCTCAGAGGTGTCTACTTGTGTGAGCCAGGGCGGTGACCCGAT 245

```

```

Db 48 SerGlyThrThrTrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGlu 67
Qy 246 GAGATCGGCTTGATGAACATCGACAGACACTCCCGTCTCTGGAG-----TACCCA--- 296
Db 68 LysCysLysArgAlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHis 87
Qy 297 -----CAGCGGGCTGGACATCAAGAACTGACCTCCCGCCCTCATCAAGAGC 350
Db 88 LysGluLysProAspLeuGluPheValLeuGluMetSerProGlnLeuIleLysThr 107
Qy 351 CACCTGCCCTACCGCTTCTGCGCTCTGACCTCCACAAATGGAGACTCCAAGGTCTAT 410
Db 108 HisLeuProSerHisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyr 127
Qy 411 ATGCTCGCAACCCCAAGATCTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGGG 470
Db 128 ValAlaArgAsnProLysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSer 147
Qy 471 ACCATGAGCTACCGAGGCACCTTCAAGAATTCGCGGAGGTTTATGAATGATAAGCTG 530
Db 148 PheMetProAspProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysVal 167
Qy 531 GGCTACGGCTCTCTGGTTTGAGCAGCTGCAGAGTTCTGGGAGCACCCGCTGAGCTCGAAC 590
Db 168 ValGlyGlySerTrpPheAspHisMetLysGlyTrpTrpAlaLysAspMetHisArg 187
Qy 591 GTGCTTTTCTCAAGTATGAAGACATCATCGGACCTGTGACGATGGTGGAGCAGCTG 650
Db 188 IleLeuTyrLeuPheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysVal 207
Qy 651 GCCAGATTCTCTGGGGTGTCTGTGACAAAGCCGCTGGAAGCCCTGACGGGACACTGC 710
Db 208 LeuGluPheLeuGluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThr 227
Qy 711 CACCAGCTGTGT-----GACCAGTGTGCAACGCTGAGGCCCTGCC----- 752
Db 228 SerPheAspValMetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIle 247
Qy 753 -----GTGGCGCGGGAAGAGTTGGGCTGTGTGGAAGACATC 788
Db 248 PheAsnHisSerIleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHis 267
Qy 789 TTCACGCTTCCATGAATGAGAAGTTTCACTGTGTATATAACAGAGATGGGAAGTGT 848
Db 268 PheThrValAlaMetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySer 287
Qy 849 GACCTCAGCTTT 860
Db 288 ThrLeuAsnPhe 291

RESULT 62
US-09-898-570-26
; Sequence 26, Application US/09898570
; Patent No. US20020123612A1
; GENERAL INFORMATION:
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
; FILE REFERENCE: 15966-776CIP
; CURRENT APPLICATION NUMBER: US/09/898,570
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/198,293
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/198,645
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/210,809
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/199,476
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,025

```

```
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/224,610
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/200,024
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/199,880
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/218,591
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/271,814
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/215,855
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 09/839,446
; PRIOR FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: POLYX
; OTHER INFORMATION: h_nh0443k08_A
US-09-898-570-26

Alignment Scores:
Pred. No.: 4,34e-24 Length: 305
Score: 396.50 Matches: 93
Percent Similarity: 50.8% Conservative: 60
Best Local Similarity: 30.9% Mismatches: 127
Query Match: 9.0% Indels: 21
DB: 3 Gaps: 5

US-10-768-158-1 (1-2419) x US-09-898-570-26 (1-305)
QY 21 ATGCGGAGAGCGAGCGGAGACCCGAGACCCCGGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaLysLeuGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
QY 78 TTCGAGTTCCATGCGTGGCGGCTCCGCGCCCTTCTGCGCGGGAAGAGATGGAGATCGCC 137
Db 21 MetGluValAlaAspGlyValProThrLeuLeuSerLysGluLysGluLysValCys 40
QY 138 AACTTCCGGTCCGCGCGGAGCGAGCTGTGGATCTACCTACCCCAAGTCCGGACGAGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTyrProLysSerGlyThrThr 60
QY 198 TTGCTGACGAGGTTGCTACTTGTGACGCGGCGTGCACCCGATGATGATCGGCTTG 257
Db 61 TrpMetHisGluLeuLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
QY 258 ATGAACATCGACGAGCAGCTCCCGTCTCTGAG-----TACCCA-----CAGCCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
QY 303 GGCTTGACATCATCAGGAACTACCTCTCCCGCTCATCAGAGCCACCTGCCCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuLeuThrHisLeuProSer 120
QY 363 CGCTTTGCGCCTCTGACCTCCACAATGGAGACTCCCAAGGTCACTATGGCTCGCAAC 422
Db 121 HisLeuLeuProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
QY 423 CCCAAGGATCTGGGTGTTCTATTATCAGTTCCACCGCTCTCTCGGACCATGAGCTAC 482
Db 141 ProlLysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSerPheMetProAsp 160
QY 483 CGAGGCACTTTCAGAAATCTCCGCGAGGTTTATGAATGATAGCTGGGCTACGGCTCC 542
Db 161 ProGlnAsnLeuGluPheTyrGluLysPheMetSerGlyLysGlyGluPheGlySer 180
QY 543 TGGTTGACGACGTCGAGGAGTTCTGCGGAGCAGCGCATGGAAGTGGCTTTTCTC 602
```

```
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200
QY 603 AAGTATGAAGACATGCGATCGGACCTGCTGACGATGCTGAGCAGCTGGCCAGATTCTGTG 662
Db 201 PheTyrGluAspIleLysGlnAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
QY 663 GGGGTGCTGTCGACAAAGGCCAGCTGGAAGCCCTGACGAGGACCTGCCACGAGCTGGTG 722
Db 221 GluLysThrTrpSerGlyAspValIleAsnLysIleValHisThrSerPheAspVal 240
QY 723 -----GACCAGTCTGCAACGCTGAGCGCTGCCCTG----- 755
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
QY 756 -----GCCCCGGGAAGATTGGGCTGTGGAGGACATCTTCCACCGTC 797
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrVal 280
QY 798 TCCATGAATGAGAAGTTTGACTTGGTGTATAAACAGAGATGGAAAGTGTGACCTCACG 857
Db 281 AlaLeuAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsn 300
QY 858 TTT 860
Db 301 Phe 301

RESULT 63
US-09-839-446-26
; Sequence 26, Application US/09839446
; Publication No. US20030050232A1
; GENERAL INFORMATION:
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: 15966-776
; CURRENT APPLICATION NUMBER: US/09/839,446
; CURRENT FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: 60/198,293
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/198,645
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/210,809
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/199,476
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,025
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/224,610
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/200,024
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/199,880
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/218,591
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/271,814
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: POLYX
; OTHER INFORMATION: h_nh0443k08_A
US-09-839-446-26

Alignment Scores:
```









```

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,065
; PRIOR FILING DATE: 2000-09-14

Alignment Scores:
Pred. No.: 4,28e-23 Length: 76
Score: 382.00 Matches: 71
Percent Similarity: 97.3% Conservative: 1
Best Local Similarity: 95.9% Mismatches: 2
Query Match: 8.7% Indels: 0
DB: 3 Gaps: 0

US-10-768-158-1 (1-2419) x US-09-989-442-149 (1-76)
QY 288 GAGTACCACAGCGCGGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAG 347
Db 1 LysfTyroGlnProGlyLeuAspIleile***GlulLeuThrSerProArgLeuileys 20
QY 348 AGCCACCTGCGCTTCTGCGCTTCTGCGCTTCCACCAATGAGACTCCAAGGTCATC 407
Db 21 SerHisLeuProTyArgPheLeuProSerAspLeuHisAenGlyAspSerIysValle 40
QY 408 TATATGCTCGCAACCCCAAGGATCTGGTGTCTTTATTATCATGTTCCACCGCTCTCTG 467
Db 41 TyrMetAlaArgAenProLysAspLeuValSerTyTyTyGlnPheHisArgSerLeu 60
QY 468 CGGACCATGAGTACCGAGGCACCTTTCAGAATTCGCGG 509
Db 61 ArgThrMetSerTyArgGly***PheGlnGluPheCysArg 74

RESULT 67
US-10-756-149-5061
; Sequence 5061, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOS
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANC
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 5061
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-5061

Alignment Scores:
Pred. No.: 9,22e-23 Length: 350
Score: 381.00 Matches: 94
Percent Similarity: 51.8% Conservative: 52
Best Local Similarity: 33.3% Mismatches: 108
Query Match: 8.7% Indels: 28
DB: 5 Gaps: 7

US-10-768-158-1 (1-2419) x US-10-756-149-5061 (1-350)
QY 54 CCGGGGAGTTCGAGAGCAAGTACTTCAGTTCATGGCGTGGCGCTGCGGCCCTTCTGCG 113
Db 13 ProGlyGlu-----TyrPheArgTyrlsGlyVal-----ProPhePro 25
QY 114 CGCGGG-----AAGATGGAGGAGATGCGC-----AACTTCCGGTGGCGGCC 155

```

```
Db 26 ValGlyLeuYrSerLeuGluSerIleSerLeuAlaGluAenThrGlnAspValArgAsp 45
Qy 156 AGCGAGGTGCGATCGTCACCTACCCCAAGTCCGCGACACGCTTCTGCGAGAGGTGTC 215
Db 46 AspAspIlePheIleIleThrTrpProLysSerGlyThrTrpMetIleGluIleIle 65
Qy 216 TACTGTGTGACCGCGCGCTGACCCCGATGAGATCGGCTTGTATGAACATCGACGAGCAG 275
Db 66 CysLeuIleLeuLysGluGlyAspProSerTrpIleArgSerValProIleTrpGluArg 85
Qy 276 CTCCCGGTCTCGGAGTACCCACAGCGCGCTGGACATCATCAAGGAACCTACCTCC 335
Db 86 AlaProTrpCysGlu---ThrIleValGlyAlaPheSerLeuProAspGlnTrpSerPro 104
Qy 336 CGCTCATCAAGACACCTCCCTACCGCTTCTGCGCTCTGACCTCCACAATGGAGAC 395
Db 105 ArgLeuMetSerSerHisLeuProIleGlnIlePheThrLysAlaPhePheSerSerLys 124
Qy 396 TCCAAGTCTATATATGCTGCAACCCCAAGATCTGGTGTCTTATATATCAAGTTC 455
Db 125 AlalysValIleIleYrMetGlyArgAsnProArgAspValValSerLeuTrpHisTrp 144
Qy 456 CACGCTCTCGCGGACCATGAGTACCGAGCACCTTTCAAGAAATTCGCGGAGGTTT 515
Db 145 SerLysIleAlaGlyGlnLeuLysAspProGlyThrProAspGlnPheLeuArgAspPhe 164
Qy 516 ATGAATGATAGTGGGTACGGCTCCTGTTTTCGACGACGTCGAGGAGTCTGGGAGCAC 575
Db 165 LeuLysGlyGluValGlnPheGlySerTrpPheAspHisIleLysGlyTrpLeuArgMet 184
Qy 576 CGCATGGAAGTCTGCAAGCTGCTTTTCTCAAGTATGAAGACATGCATCGGACCTGGTGA 635
Db 185 LysGlyLysAspAsnPheLeuPheIleThrTrpGluGluLeuGlnGlnAspLeuGlnGly 204
Qy 636 ATGTGTGAGCAGCTGGCCAGATTCCTGGGGTGCTGTGTGACAGGCCACCTGGAAGCC 695
Db 205 SerValGluArgIleCysGlyPheLeuGlyArgProLeuGlyLysGluAlaLeuGlySer 224
Qy 696 CTGACGGAGCAC-----TGCACACGCTGTGTGGACCAAGTGTGCAACGCTGAGGCC 746
Db 225 ValValAlaHisSerThrPheSerAlaMetLysAlaAenThrMetSerAsnTrpThrLeu 244
Qy 747 CTGCCC-----GTGGCGCGGGGAAGAGTGGG 773
Db 245 LeuProProSerLeuLeuAspHisArgGlyAlaPheLeuArgLysGlyValCysGly 264
Qy 774 CTGTGGAGGACATCTTCACCGTCTCCATGAATGAGAGTTTGACTGGTGTATAACAGC 833
Db 265 AspTrpLysAsnHisPheThrValAlaGlnSerGluAlaPheAspArgAlaTrpArgLys 284
Qy 834 AAGATG 839
Db 285 GlnMet 286
```

## RESULT 68

```
US-11-097-143-11343
; Sequence 11343, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CLO00728
; CURRENT FILING DATE: 2005-04-04
; PRIOR FILING DATE: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
```

```
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11343
; LENGTH: 338
; TYPE: PRN
; ORGANISM: DROSOPHILA
US-11-097-143-11343
```

## Alignment Scores:

Pred. No.:	2,7e-21	Length:	338
Score:	363.50	Matches:	91
Percent Similarity:	48.3%	Conservative:	53
Best Local Similarity:	30.5%	Mismatches:	119
Query Match:	8.3%	Indels:	35
DB:	6	Gaps:	7

US-10-768-158-1 (1-2419) x US-11-097-143-11343 (1-338)

```
Qy 12 GGCGCGCGCTGCGGAGAGGCGGAGAGCCCGACACCCCGACCCCG-----GGGAGTTC 65
Db 23 GlyGlyTrpSerSerIlePheAlaSerSerLysProSerValProValValGlyAsnTrp 42
Qy 66 GAGAGCAAGTACTTCGAGTTCCTGCGGTGCGCGCTCGCGCTTCTGCGCGGGAAGATG 125
Db 43 GluGlnArgPhe-----CysArgLeuAlaAspThrPheGlnProValLeu 57
Qy 126 GAGGAGATCCGCCAACTTCCCGGTGCGCGCCCGAGCGAGCTGTGGATCGTCACCTACCCCAAG 185
Db 58 AspArgValTrpAspPheGluValArgAspAspValTrpIleValThrLeuProLys 77
Qy 186 TCGGCGACCAAGTGTGTCGAGGAGGTGCTACTTGTGTGAGCCAGGGCGCTGACCCCGAT 245
Db 78 CysGlyThrTrpMetGlnGluLeuAlaTrpLeuValIleAsnGluCysAspPheGlu 97
Qy 246 GAGATCGGCTTGATGAACATCGACGACGAGCTCCCGTCTCTGGAGTAC----- 293
Db 98 ThrAlaLysSerValAspLeuThrHisArgSerProPheLeuGluPheAsnGlyValVal 117
Qy 294 -----CCACAGCGCGGCGCTTGACATCATCAAGAACTGACCTCTCCCGCGCTCATC 344
Db 118 ProAsnValProHisAspThrIleAlaAlaAsnAlaLeuProSerProArgLeuIle 137
Qy 345 AAGACCCACCTGCGCTACCGCTTCTGCGCTCTGACCTCCACAATGGAGACTCCAAGTC 404
Db 138 LysSerHisLeuProAlaTrpMetLeuProArgGlnIleTrpSerLysArgProLysIle 157
Qy 405 ATCTATATGCTCGCAACCCCAAGGATCTGGTGTGTCTTATTATTAGTTCCACCGCTCT 464
Db 158 IleTrpValTrpArgAsnProLysAspAlaIleSerTrpPheHisHisTrpArgGly 177
Qy 465 CTGCGGACCATGAGTACCGGACCTTTCAAGATTTCTGCCGAGAGTTTATGAATGAT 524
Db 178 Met-----ValGlyTrpGlnGlyThrLysSerAspPheMetHisSerPheIleAspGly 195
Qy 525 AAGCTGGCTACGGCTCTCTGGTTTGACAGCTGTCAGGAGTTCGGGAGCACCGCATGAC 584
Db 196 TyrValAsnPheThrProCysTrpProHisIleLeuAspPheTrpGlnLeuArgHisGlu 215
Qy 585 TCGAACGCTGCTTTTCTCAAGTATGAAGACATGATCGGACCTGGTGAGCATGCTGGAG 644
Db 216 ProAsnIlePhePheThrSerTrpGluArgMetLysGlyGlnLeuGlyGlnValIleSer 235
```

```
Qy 645 CAGTGGCCGATCTCTGGGGGTCTCTGTGACAAAGCCAGCTGGAAGCCCTCACCGAG 704
      :::::::::::::::::::: ||| ::::::::::::::::::::
Db 236 GluValAlaGlnPheLeuGluArgSerValSerGlnGlnMetGlnMetGlnArg 255
      :::::::::::::::::::: ||| ::::::::::::::::::::
Qy 705 CACTGC-----CACCAGCTGGTGGACAG---TGCTGCAAC----- 737
      ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 256 HisLeuSerPheGluSerMetArgAspAsnProAlaCysAsnHisValLysGluPheGlu 275
      ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 738 -----CCTGAGGCCCTGCCCTGGCGGGCGGGGGAAGA 767
      ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 276 SerMetLysAlaAlaGluArgGluValGluGluPheArgPheValArgArgGlyVal 295
      ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 768 GTTCGGCTGTGGAAGGACATCTTCACGCTCTCCATGAATGAGAAAGTTTGACTTG 821
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 296 ValGlySerHisLysAspGluLeuThrAlaAspIlelleArgGluPheAspLeu 313
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 69
US-11-097-143-14628
; Sequence 14628, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14628
; LENGTH: 346
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-14628

Alignment Scores:
Pred. No.: 5,92e-21 Length: 346
Score: 259.50 Matches: 82
Percent Similarity: 46.1% Conservative: 35
Best Local Similarity: 32.3% Mismatches: 88
Query Match: 8.2% Indels: 49
DB: 6 Gaps: 5

US-10-768-158-1 (1-2419) x US-11-097-143-14628 (1-346)

Qy 84 TTCATGCGCTGCGG-----CTGCGCGCCC 107
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 27 PheHisGlyGluArgThrGlyPheValGlnValGlySerGluGlyTyPhePheProHis 46
      ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 108 TTCTGCCCGGAAGATGAGAGATCCCACTTCCCGTCCGCCCGCCAGCGACGTGTGG 167
      ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 47 LysTyLysAspGluAlaGluArgTyTyAsnPheGluAlaArgProAspValTrp 66
      ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 168 ATGCTCACTACCCCAAGTCCCGCACCGACCTTGCTGCGAGGAGTGGTCTACTTGGTGAGC 227
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 67 IleAlaThrValProArgSerGlyThrTrpThrGlnGluLeuIleTrpLeuValAla 86
Qy 228 CAGGGCCCTCACCCTGATGAGATCGGCTTGTATGACATCGACGAGCTCCCGTCTCTG 287
      ||| ||| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 87 AsnGlyLeuAspPheGluHisAlaGlnGluArgProLeuThrGluArgPheProPhe 106
      ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 288 GAGTACCACA----- 296
      ||| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 107 GluPheProLeuPheValHisProLysIleLysGluGlnGluGluAsnArgAsp 126
Qy 297 -----CAGCGCGCGCTGGACATCATCATCAAG 320
      ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 127 SerAlaGluAlaLeuGluPheIleGluLysIleAlaArgProGlyTyArgGluAlaLeuSer 146
Qy 321 GAACGT-----ACCTCTCCCGCTCATCAAGAGCCACCTGCGCTACCGCTTCTGCC 374
      ||| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 147 GluIleProArgSerGlnArgPheIleLysThrHisPheProSerLeuMetPro 166
Qy 375 TCTGACCTCCACAATGGAGACTCCCAAGTCTATCTATATGCTCGCAACCCCAAGATCTG 434
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 ProSerValLeuGluLysLysCysLysValIleTyValValArgAspProLysAspVal 186
Qy 435 GTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGCTACCGAGGCACCTTT 494
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 AlaValSerTyTyHisLeuAsnArgLeuPheArgThrGlnGlyTyValGlyAspPhe 206
Qy 495 CAAGAATCTCGCGAGGTTTATGAATGATAAGCTGGCTACGGCTCTCTGG----- 545
      ::| ::| ::| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 207 GluArgTyTrpHisTyPheGlnAsn-----GlyLeuAsnProTrpLeuProTy 223
Qy 546 TTTGAGCAGCTGCAGGAGTTCTGGGAGCAGCGCATCGACTCGAACGTCTTTTCTCAAG 605
      ::| ::| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 TySerHisValLysGluAlaArgGluHisAlaHisLeuSerAsnValLeuPheLeuArg 243
Qy 606 TATGAACATATGATCGGACCTGTCACCGATGTCAGCATGTCGAGCAGCTGGCAGATTCTCTGGGG 665
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 244 TyGluAspMetLeuAlaAspLeuProGlyAlaIleAsnSerIleAlaSerPheLeuGlu 263
Qy 666 GTGTCCTGTGACAGGCCCGCTGGAAGCCCTGACGGAGCAGC 707
      ::| ::| ::| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 264 CysProLysProGluAspMetAspArgLeuLeuAspHis 277

RESULT 70
US-11-152-635-11
; Sequence 11, Application US/11152635
; Publication No. US20060024708A1
; GENERAL INFORMATION:
; APPLICANT: Squires, James
; TITLE OF INVENTION: Porcine Sulfotransferase 2A1 Polynucleotide Sequence, Protein, an
; TITLE OF INVENTION: Methods of Use for Same
; FILE REFERENCE: P06815US01
; CURRENT APPLICATION NUMBER: US/11/152,635
; CURRENT FILING DATE: 2005-06-14
; PRIOR APPLICATION NUMBER: 60/580,540
; PRIOR FILING DATE: 2005-06-17
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-152-635-11

Alignment Scores:
Pred. No.: 9.84e-21 Length: 285
Score: 356.50 Matches: 83
Percent Similarity: 53.0% Conservative: 59
Best Local Similarity: 31.0% Mismatches: 109
Query Match: 8.1% Indels: 17
DB: 6 Gaps: 5

US-10-768-158-1 (1-2419) x US-11-152-635-11 (1-285)
```







```
Db 109 PhePheLysSerLysAlaLysMetIleTyrIleIleArgAsnProArgAspIleVal 128
Qy 441 TCTTATTATGTTCCACCGCTCTCTCGGACCATGAGTACCGAGGACACCTTTCAAGAA 500
Db 129 SerGlyTyrPhePheTrpLysSerThrAsnLeuValLysArgProGluSerLeuGlu 148
Qy 501 TTCTGCGGAGGTTTATGATGATAAGCTGGGCTACCGCTCTCTGTTTGACACAGTGCAG 560
Db 149 TyrPheGluTrpPheIleGlnGlyAsnValProTyrGlySerTrpPheAspHisIleArg 168
Qy 561 GAGTCTCTGGGAGCACCGCATGGATCTGCAAGCTGCTTTTCTCAAGTATGAGAGCATGCAT 620
Db 169 GlyTrpLeuProMetArgAspLysGluAsnValLeuLeuLeuSerTyrGluGluLeuLys 198
Qy 621 CGGACCTGTTGACGATGGTGGAGCAGCTGCCAGATTCCTCGGGGGTCTCTGCGACAAG 680
Db 189 ArgAspThrArgSerAlaValGluLysIleCysGlnPheLeuGlyLysLeuGluPro 208
Qy 681 GCCCAGCTGGAGCCCTGACGAGCAGTGCACCGAGCTGGTGGACCATGCTGCTGCAAC--- 737
Db 209 GluGluLeuSerSerValValGluAsnSerSerPheGlnValMetLysGluAsnMet 228
Qy 738 -----GCTGAGGCCCTGCCCTGGGC-----CGG 761
Db 229 SerAsnPheSerLeuLeuLysGlyLeuHisLeuGlyAspThrGlyCysLeuLeuArgLys 248
Qy 762 GGAAGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCCATGATGAATGAACTTTGACTTG 821
Db 249 GlyThrProGlyAspTrpLysAsnTyrPheThrValAlaGlnAlaGluAlaPheAspLys 268

RESULT 74
US-11-152-635-10
; Sequence 10, Application US/11152635
; Publication No. US20060024708A1
; GENERAL INFORMATION:
; APPLICANT: Squires, James
; TITLE OF INVENTION: Porcine Sulfotransferase 2A1 Polynucleotide Sequence, Protein, an
; FILE OF INVENTION: Methods of Use for Same
; FILE REFERENCE: P06815US01
; CURRENT APPLICATION NUMBER: US/11/152,635
; CURRENT FILING DATE: 2005-06-14
; PRIOR FILING DATE: 2005-06-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Porcine
US-11-152-635-10

Alignment Scores:
Pred. No.: 5.63e-20 Length: 285
Score: 347.50 Matches: 88
Percent Similarity: 50.3% Conservative: 56
Best Local Similarity: 30.8% Mismatches: 117
Query Match: 7.9% Indels: 25
DB: 6 Gaps: 6

US-10-768-158-1 (1-2419) x US-11-152-635-10 (1-285)
Qy 21 ATGCGGAGAGCGAGCGCGAGACCCCGGAGTTCGAGCAAGTACTTC 80
Db 1 MetThrGluGluGluVal-----ArgPheGluGlyIlePhePhe 13
Qy 81 GAGTTCATCGCTCGGCTCCGCCCTTCTGCGCGGAGAGATGGAGGATCGCCAAAC 140
Db 1 MetThrGluGluGluVal-----ArgPheGluGlyIlePhePhe 13
Qy 81 GAGTTCATCGCTCGGCTCCGCCCTTCTGCGCGGAGAGATGGAGGATCGCCAAAC 140
Db 1 MetThrGluGluGluVal-----ArgPheGluGlyIlePhePhe 13
```

```
Db 14 Pro---LysGlnIleLeuSerProGluMetLeuGlnGluValArgGluGlu----- 29
Qy 141 TTCCCGGTGGCGCCAGCAGCTGTGTGATCGTCACTTACCCCAAGTCCCGGACAGCTTG 200
Db 30 PheThrPheLysGluGluAspValLeuIleLeuThrPheProLysSerGlyThrAsnTrp 49
Qy 201 CTGCAGAGGTGTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTCAATG 260
Db 50 MetIleGluIleLeuLysLeuIleLeuSerLysGlyAspThrLysTrpIleGlnSerVal 69
Qy 261 AACATCAGCAGCAGCTCCCGCTCTCGGAGTACCCACAGCCGGCCCTGGACATCAAG 320
Db 70 ProAsnTrpAspArgSerProTrpLeuGlu---SerIleSerGlyTyrGluAsnLeuLys 88
Qy 321 GAACTGACCTCTCCCGCTCATCAAGAGCAGCTGCCCTACCGCTTCTGCTCCCTGAC 380
Db 89 AsnLysGluGlyProArgLeuIleSerSerHisLeuProIleGlnPheProLysAla 108
Qy 381 CTCCACAATGGAGACTCCAAGTCACTATATGGCTCGCAACCCCAAGGATCTGGTGTG 440
Db 109 PhePheLysSerLysAlaLysMetIleTyrIleIleArgAsnProArgAspIleVal 128
Qy 441 TCTTATTATCAGTTCACCGCTCTCTCGGAGCAGTACCGAGGACACCTTTCAAGAA 500
Db 129 SerGlyTyrPhePheTrpLysSerThrAsnLeuValLysArgProGluSerLeuGlu 148
Qy 501 TTCTGCGGAGGTTTATGATGATAAGCTGGGCTACCGCTCTCTGTTTGACACAGTGCAG 560
Db 149 TyrPheGluTrpPheIleGlnGlyAsnValProTyrGlySerTrpPheAspHisIleArg 168
Qy 561 GAGTCTCTGGGAGCACCGCATGGACTCGAACGTGCTTTTCTCAAGTATGAGAGCATGCAT 620
Db 169 GlyTrpLeuProMetArgAspLysGluAsnValLeuIleLeuSerTyrGluGluLys 188
Qy 621 CGGACCTGTTGACGATGGTGGAGCAGCTGCCAGATTCCTCGGGGGTCTCTGTCGACAAG 680
Db 189 ArgAspThrArgSerAlaValGluLysIleCysGlnPheLeuGlyLysLeuGluPro 208
Qy 681 GCCCAGCTGGAGCCCTGACGAGCAGTGCACCGAGCTGGTGGACCATGCTGCTGCAAC--- 737
Db 209 GluGluLeuSerSerValValGluAsnSerSerPheGlnValMetLysGluAsnMet 228
Qy 738 -----GCTGAGGCCCTGCCCTGGGC-----CGG 761
Db 229 SerAsnPheSerLeuLeuLysGlyLeuHisLeuGlyAspThrGlyCysLeuLeuArgLys 248
Qy 762 GGAAGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCCATGATGAATGAACTTTGACTTG 821
Db 249 GlyThrProGlyAspTrpLysAsnTyrPheThrValAlaGlnAlaGluAlaPheAspLys 268

822 GTGTATAAACAGAGATG 839
269 LeupheGlnGluLysMet 274

RESULT 75
US-09-795-926-2
; Sequence 2, Application US/09795926
; Patent No. US20020098486A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Brin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
```

```
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 303
; TYPE: PRT
; ORGANISM: homo sapiens
; ORGANISM: homo sapiens
US-09-795-926-2

Alignment Scores:
Pred. No.:      2,91e-16      Length:      303
Score:          303.50        Matches:     74
Percent Similarity: 48.7%      Conservative: 61
Best Local Similarity: 26.7%    Mismatches:  99
Query Match:     6.9%         Indels:      43
DB:              3           Gaps:           9

US-10-768-158-1 (1-2419) x US-09-795-926-2 (1-303)
Qy 75 TACTTCGAGTTCCATGCGTCCGCTCCGCGC---CCCTTCTGCCCGGGAAG---ATGGAG 128
Db 27 PhePheThrTyrGlnGlyIleProTyrProIleThrMetCysThrSerGluThrPheGln 46
Qy 129 GAGATCGCCCAACTTCCCGGTCCGCGCCAGCAGCTGGATCGTCACCTACCCCAAGTCC 188
Db 47 AlaLeuAspThrPheGluAlaArgHisAspAspIleValLeuAlaSerTyrProLysCys 66
Qy 189 GGCACCAAGC-----TTGTCGACGAGTGGTCTACTTTGGTGAGCCAGGCGCT 236
Db 67 GlySerAsnTrpIleLeuHisIleValSerGluLeuIleTyrAlaValSerLys----- 84
Qy 237 GACCCCGATGATCGCGCTTGATGAACATCGACGAGCTCCGCGTCTGGAGTACCCA 296
Db 85 -----LysLysTyrLysTyrPro 90
Qy 297 CAG---CCGGCGCTGGACATC-----ATCAGGAAGTCAAC 329
Db 91 GluPheProValLeuGluCysGlyAspSerGluLysTyrGlnArgMetLysGlyPhePro 110
Qy 330 TCTCCCGCGCTCATCAAGAGCCACCTCCCTACCGCTTCTGCGCTCTGACCTCCACAAT 389
Db 111 SerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePheGlu 130
Qy 390 GGAGACTCCAAAGTCACTATATGGTCGCGCAACCCCAAGGATCTGGTGTCTTATTAT 449
Db 131 AsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPheLeu 150
Qy 450 CAGTTCCACCGCTCTTCGCGCAACCATGAGTACCGAGCACCCTTCAAGATTTCTGCGG 509
Db 151 HisPheHisAsnAspValProAspIleProSerTyrGlySerTyrAspGluPheArg 170
Qy 510 AGGTTTATGAATATAAGCTGGGTACGGCTCCCTGGTTTGGACGCGTGGAGGATCTTGG 569
Db 171 GlnPheMetLysGlyGlnValSerTrpGlyArgTyrPheAspPheAlaIleAsnTrpAsn 190
Qy 570 GAGACCGCATGAGTCAAGCTGCTTTCTCAAGTATGAAGACATGATCGGAGCTG 629
Db 191 LysHisLeuAspGlyAspAsnValLysPheIleLeuTyrGluAspLeuLysGluAsnLeu 210
Qy 630 GTGACGATGGTGGAGCAGCTGGCAGATTCTCTGGGGGTGCTCTGTGACAGGCCAGCTG 689
Db 211 AlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPhePheLeuThrGlyGluGlnIle 230
Qy 690 GAAGCCCTGACGGAGCACTGCCACCATGCTGGTGGAGCAGCTGCTGCAACGCTGAGGCCCTG 749
Db 90 -----LysLysTyrLysTyrPro 90

Db 231 GlnThrIleSerVal-----GlnSerThrPheGlnAlaMetArgAlaLysSerGln 247
Qy 750 -----CCGTFGGGC-----CGGGGAAGAGTTGGCTCTTGAAG 782
Db 248 AspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyGluValGlyAspTrpLys 267
Qy 783 GACATCTTCCCGCTCTCCATGAATGAGAAGTTTGACTTGGTGATATAACAG 833
Db 268 AsnLeuPheSerGluIleGlnAsnGlnMetAspGluLysPheLysGlu 284

RESULT 76
US-10-364-774-2
; Sequence 2, Application US/10364774
; Publication No. US2003014497A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/10/364,774
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 303
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774-2

Alignment Scores:
Pred. No.:      2,91e-16      Length:      303
Score:          303.50        Matches:     74
Percent Similarity: 48.7%      Conservative: 61
Best Local Similarity: 26.7%    Mismatches:  99
Query Match:     6.9%         Indels:      43
DB:              4           Gaps:           9

US-10-768-158-1 (1-2419) x US-10-364-774-2 (1-303)
Qy 75 TACTTCGAGTTCCATGCGTCCGCTCCGCGC---CCCTTCTGCCCGGGAAG---ATGGAG 128
Db 27 PhePheThrTyrGlnGlyIleProTyrProIleThrMetCysThrSerGluThrPheGln 46
Qy 129 GAGATCGCCCAACTTCCCGGTCCGCGCCAGCAGCTGGATCGTCACCTACCCCAAGTCC 188
Db 47 AlaLeuAspThrPheGluAlaArgHisAspAspIleValLeuAlaSerTyrProLysCys 66
Qy 189 GGCACCAAGC-----TTGTCGACGAGTGGTCTACTTTGGTGAGCCAGGCGCT 236
Db 67 GlySerAsnTrpIleLeuHisIleValSerGluLeuIleTyrAlaValSerLys----- 84
Qy 237 GACCCCGATGATCGCGCTTGATGAACATCGACGAGCTCCGCGTCTGGAGTACCCA 296
Db 85 -----LysLysTyrLysTyrPro 90
Qy 297 CAG---CCGGCGCTGGACATC-----ATCAGGAAGTCAAC 329
Db 91 GluPheProValLeuGluCysGlyAspSerGluLysTyrGlnArgMetLysGlyPhePro 110
Qy 330 TCTCCCGCGCTCATCAAGAGCCACCTCCCTACCGCTTCTGCGCTCTGACCTCCACAAT 389
Db 111 SerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePheGlu 130
Qy 390 GGAGACTCCAAAGTCACTATATGGTCGCGCAACCCCAAGGATCTGGTGTCTTATTAT 449
Db 131 AsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPheLeu 150
Qy 450 CAGTTCCACCGCTCTTCGCGCAACCATGAGTACCGAGCACCCTTCAAGATTTCTGCGG 509
Db 151 HisPheHisAsnAspValProAspIleProSerTyrGlySerTyrAspGluPheArg 170
Qy 510 AGGTTTATGAATATAAGCTGGGTACGGCTCCCTGGTTTGGACGCGTGGAGGATCTTGG 569
Db 171 GlnPheMetLysGlyGlnValSerTrpGlyArgTyrPheAspPheAlaIleAsnTrpAsn 190
Qy 570 GAGACCGCATGAGTCAAGCTGCTTTCTCAAGTATGAAGACATGATCGGAGCTG 629
Db 191 LysHisLeuAspGlyAspAsnValLysPheIleLeuTyrGluAspLeuLysGluAsnLeu 210
Qy 630 GTGACGATGGTGGAGCAGCTGGCAGATTCTCTGGGGGTGCTCTGTGACAGGCCAGCTG 689
Db 211 AlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPhePheLeuThrGlyGluGlnIle 230
Qy 690 GAAGCCCTGACGGAGCACTGCCACCATGCTGGTGGAGCAGCTGCTGCAACGCTGAGGCCCTG 749
Db 90 -----LysLysTyrLysTyrPro 90
```

QY 297 CAG-----CCGGCGCTGGACATC-----ATCAAGGAACGTGACC 329  
Db 91 GluPheProValLeuGluCysGlyAspSerGluLysTyrGlnArgMetLysGlyPhePro 110  
QY 330 TCTCCCGCCCTCATCAAGAGCCACCTGCCCTACCGCTTCTGCTCTGACCTCCACCAAT 389  
Db 111 SerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePheGlu 130  
QY 390 GGAGACTCCAAGGTCTATATGCTCGCAACCCCAAGGATCTGGTGTCTTATTAT 449  
Db 131 AsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPheLeu 150  
QY 450 CAGTTCCACCGCTCTCTCGGGACCACTGACCTACCGAGCACCTTTCAAGAATTTGCCGG 509  
Db 151 HisPheHisAsnAspValProAspIleProSerTyrGlySerTyrAspGluPhePheArg 170  
QY 510 AGGTTTATGAATGATAAGCTGGCTACCGCTCCTCGTTTGAGCAGCTGCAGGAGTTCTGG 569  
Db 171 GlnPheMetLysGlyGlnValSerTrpGlyArgTyrPheAspPheAlaIleAsnTrpAsn 190  
QY 570 GAGCACCCTGAGCTGCAACCGTCTTTTCTCAAGTATGAAGACATGCTCGGACCTG 629  
Db 191 LysHisLeuAspGlyAspAsnValLysPheIleLeuTyrGluAspLeuLysGluAsnLeu 210  
QY 630 GTGACGATGGTGGAGCAGCTGCCAGATTCCTGGGGGTGCTGTGTGCAAGGCCCGAGCTG 689  
Db 211 AlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPhePheLeuThrGlyGluGlnIle 230  
QY 690 GAAGCCCTGCGGAGCAGCTGCCACAGCTGGTGGACCACTGCTGCAACGCTGAGGCCCTG 749  
Db 231 GlnThrIleSerVal-----GlnSerThrPheGlnAlaMetArgAlaLysSerGln 247  
QY 750 -----CCGCTGGGCG-----CGGGGAAGAGTTGGGCTGTGGGAG 782  
Db 248 AspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyGluValGlyAspTrpLys 267  
QY 783 GACATCTTCACGCTCCCATGAATGAGAAAGTTTGACTTGGTGTATPAAACAG 833  
Db 268 AsnLeuPheSerGluIleGlnAsnGlnGluMetAspGluLysPheLysGlu 284

RESULT 77

US-10-072-012-664  
; Sequence 664, Application US/10072012  
; Publication No. US20040033493A1  
; GENERAL INFORMATION:  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Zernusen, Bryan  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Li, Li  
; APPLICANT: Gangolli, Esha  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Anderson, David W.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Taupier Jr, Raymond J.  
; APPLICANT: Gusev, Vladimir Y.  
; APPLICANT: Colman, Steven D.  
; APPLICANT: Wolenc, Adam R.  
; APPLICANT: Pena, Carol E. A  
; APPLICANT: Furtak, Katarzyna  
; APPLICANT: Grosse, William M.  
; APPLICANT: Alsbrook II, John P.  
; APPLICANT: Lepley, Denise M.  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Burgess, Catherine E.  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-258  
; CURRENT APPLICATION NUMBER: US/10/072,012  
; CURRENT FILING DATE: 2002-01-31

; PRIOR APPLICATION NUMBER: 60/265,102  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 60/265,514  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,517  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,412  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,395  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/266,406  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 60/266,767  
; PRIOR FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: 60/267,057  
; PRIOR FILING DATE: 2001-02-07  
; PRIOR APPLICATION NUMBER: 60/266,975  
; PRIOR FILING DATE: 2001-02-07  
; PRIOR APPLICATION NUMBER: 60/267,459  
; PRIOR FILING DATE: 2001-02-08  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1391  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 664  
; LENGTH: 303  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-072-012-664  
  
Alignment Scores:  
Pred. No.: 2,91e-16 Length: 303  
Score: 303.50 Matches: 74  
Percent Similarity: 48.7% Conservative: 61  
Best Local Similarity: 26.7% Mismatches: 99  
Query Match: 6.9% Indels: 43  
DB: 4 Gaps: 9  
  
US-10-768-158-1 (1-2419) x US-10-072-012-664 (1-303)  
QY 75 TACTTCGAGTTCCATGCGGTGCGGTGCGG---CCCTTCGCGCGGGAAG---ATGGAG 128  
Db 27 PhePheThrTyrGlnGlyIleProTyrProIleThrMetCysThrSerGluThrPheGln 46  
QY 129 GAGATCGCCCACTTCCCGGTGCGGCCCGCCAGCAGCGTGGATCGTCACCTACCCCAAGTCC 188  
Db 47 AlaLeuAspThrPheGluAlaArgHisAspAspIleValLeuAlaSerTyrProLysCys 66  
QY 189 GGCACACAGC-----TTGCTGACGAGGTGGTCTACTTGGTGAGCCAGGCGCT 236  
Db 67 GlySerAsnTrpIleLeuHisIleValSerGluLeuIleTyrAlaValSerLys----- 84  
QY 237 GACCCCGATGAGATCGGCTTGATGAACATCAGCAGCAGCTCCCGGTCTCTGGAGTACCA 296  
Db 85 -----LysLysTyrLysTyrPro 90  
QY 297 CAG---CCGGCGCTGGACATC-----ATCAAGGAACGTGACC 329  
Db 91 GluPheProValLeuGluCysGlyAspSerGluLysTyrGlnArgMetLysGlyPhePro 110  
QY 330 TCTCCCGCCCTCATCAAGAGCCACCTGCCCTACCGCTTCTGCTCTGACCTCCACCAAT 389  
Db 111 SerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePheGlu 130  
QY 390 GGAGACTCCAAGGTCTATATGCTCGCAACCCCAAGGATCTGGTGTCTTATTAT 449  
Db 131 AsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPheLeu 150  
QY 450 CAGTTCCACCGCTCTCTCGGGACCACTGACCTACCGAGCACCTTTCAAGAATTTGCCGG 509  
Db 151 HisPheHisAsnAspValProAspIleProSerTyrGlySerTyrAspGluPhePheArg 170  
QY 510 AGGTTTATGAATGATAAGCTGGCTACCGCTCCTCGTTTGAGCAGCTGCAGGAGTTCTGG 569

Db 171 GlnPheMetLysGlyGlnValSerTrpGlyArgTyrPheAspPheAlaIleAsnTrpAsn 190  
Qy 570 GAGCACCAGCTGGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATGCATCGGGACCTG 629  
Db 191 LysHisLeuAspGlyAspAsnValLysPheIleLeuTyrGluAspLeuLysGluAsnLeu 210  
Qy 630 GTGACGATGGTGAGCAGCTGGCCAGATTCTCTGGGGGTGCTCTGTGACAAGGCCAGCTG 689  
Db 211 AlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPhePheLeuThrGlyGluGlnIle 230  
Qy 690 GAAGCCCTGACGGAGCACTGCCACAGCTGTGGACCAAGTCTGCAACCGCTGAGGCCCTG 749  
Db 231 GlnThrIleSerVal-----GlnSerThrPheGlnAlaMetArgAlaLysSerGln 247  
Qy 750 -----CCCGTGGGC-----CGGGGAAGATTGGGCTGTGGAAG 782  
Db 248 AspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyGluValGlyAspTrpLys 267  
Qy 783 GACATCTTCACCGCTCCCATCAATGAGAAGTTTTCACCTTGGTGTATAAACAG 833  
Db 268 AsnLeuPheSerGluIleGlnAsnGlnGluMetAspGluLysPheLysGlu 284

RESULT 78

US-10-468-125-3  
; Sequence 3, Application US/10468125  
; Publication No. US20040082061A1  
; GENERAL INFORMATION:  
; APPLICANT: ASTROMOFF, Anna  
; APPLICANT: AU-YOUNG, Janice  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: DING, Li  
; APPLICANT: DUGAN, Brendan M.  
; APPLICANT: FORSYTHE, Ian J.  
; APPLICANT: GIETZEN, Kimberly J.  
; APPLICANT: GRIFFIN, Jennifer A.  
; APPLICANT: LEE, Ernestine A.  
; APPLICANT: LU, Yan  
; APPLICANT: RICHARDSON, Thomas W.  
; APPLICANT: RING, Huijun Z.  
; APPLICANT: SANJANWALA, Madhusudan  
; APPLICANT: SWARNAKAR, Anita  
; APPLICANT: CHAWLA, Narinder K.  
; APPLICANT: WARREN, Bridget A.  
; APPLICANT: XU, Yuming  
; APPLICANT: YUE, Henry  
; APPLICANT: ZEBARJADIAN, Yeganeh  
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES  
; FILE REFERENCE: PI-0363 USN  
; CURRENT APPLICATION NUMBER: US/10/468,125  
; PRIOR FILING DATE: 2003-08-15  
; PRIOR APPLICATION NUMBER: PCT/US02/04918  
; PRIOR FILING DATE: 2002-02-14  
; PRIOR APPLICATION NUMBER: US 60/269,643  
; PRIOR FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/271,332  
; PRIOR FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: US 60/276,767  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: US 60/282,077  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: US 60/285,447  
; PRIOR FILING DATE: 2001-04-19  
; PRIOR APPLICATION NUMBER: US 60/287,060  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/288,543  
; PRIOR FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PERL Program  
; SEQ ID NO 3  
; LENGTH: 303  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:

; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: 7491172CD1  
US-10-468-125-3  
Alignment Scores: 2.91e-16 Length: 303  
Pred. No.: 303.50 Matches: 74  
Score: 48.7% Conservative: 61  
Best Local Similarity: 26.7% Mismatches: 99  
Query Match: 6.9% Indels: 43  
DB: 4 Gaps: 9  
US-10-768-158-1 (1-2419) x US-10-468-125-3 (1-303)  
Qy 75 TACTTCAGTTCATGCGGTGCGGTCCG---CCCTTCTCGCGCGGAAG---ATGAG 128  
Db 27 PhePheThrTyrGlnGlyIleProTyrProIleThrMetCysThrSerGluThrPheGln 46  
Qy 129 GAGATCGCAACTTCCCGGTGCGGCCCGCAGCGATGTGGATCGTCACCTACCCCAAGTCC 188  
Db 47 AlaLeuAspThrPheGluAlaArgHisAspIleValLeuAlaSerTyrProLysCys 66  
Qy 189 GGCACACAGC-----TTGCTCAGAGGTGCTACTTGGTGAGCCAGGGCGCT 236  
Db 67 GlySerAsnTrpIleLeuHisIleValSerGluLeuIleTyrAlaValSerLys----- 84  
Qy 237 GACCCGATGATCGCTTGATGAACATCGACGAGCAGCTCCCGGTCTCTGGAGTACCCA 296  
Db 85 -----LysLysTyrLysTyrPro 90  
Qy 297 CAG---CGGGCCTGGACATC-----ATCAAGGAACATGACC 329  
Db 91 GluPheProValLeuGluCysGlyAspSerGluLysTyrGlnArgMetLysGlyPhePro 110  
Qy 330 TCTCCCGCTCATCAAGAGCCACTGCCCTTACCGCTTTCGCGCTCTGACCTCCACAT 389  
Db 111 SerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePheGlu 130  
Qy 390 GGAGACTCCAGGTCTATATATGCTCGCAGCCCAAGGATCTGGTGGTCTTATAT 449  
Db 131 AsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPheLeu 150  
Qy 450 CAGTTCACCGCTCTCTCGCGACCATGAGCTACCGAGCACCTTTCAAGAAATTCTGCCGG 509  
Db 151 HisPheHisAsnAspValProAspIleProSerTyrGlySerTrpAspGluPheArg 170  
Qy 510 AGGTTTATGAATGAAGCTGGGTACGGCTCTCGTTTGGACGACGCTGACGAGTCTTGG 569  
Db 171 GlnPheMetLysGlyGlnValSerTrpGlyArgTyrPheAspPheAlaIleAsnTrpAsn 190  
Qy 570 GAGCACCAGCTGGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATGCATCGGGACCTG 629  
Db 191 LysHisLeuAspGlyAspAsnValLysPheIleLeuTyrGluAspLeuLysGluAsnLeu 210  
Qy 630 GTGACGATGGTGAGCAGCTGGCCAGATTCTCTGGGGGTGCTCTGTGACAAGGCCAGCTG 689  
Db 211 AlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPhePheLeuThrGlyGluGlnIle 230  
Qy 690 GAAGCCCTGACGGAGCACTGCCACAGCTGTGGACCAAGTCTGCAACCGCTGAGGCCCTG 749  
Db 231 GlnThrIleSerVal-----GlnSerThrPheGlnAlaMetArgAlaLysSerGln 247  
Qy 750 -----CCCGTGGGC-----CGGGGAAGATTGGGCTGTGGAAG 782  
Db 248 AspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyGluValGlyAspTrpLys 267  
Qy 783 GACATCTTCACCGCTCCCATCAATGAGAAGTTTTCACCTTGGTGTATAAACAG 833  
Db 268 AsnLeuPheSerGluIleGlnAsnGlnGluMetAspGluLysPheLysGlu 284  
RESULT 79  
US-11-134-241-2  
; Sequence 2, Application US/11134241



```
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USN 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: USN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 150
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-900A-150

Alignment Scores:
Pred. No.: 4,67e-16 Length: 295
Score: 301.00 Matches: 85
Percent Similarity: 47.8% Conservative: 58
Best Local Similarity: 28.4% Mismatches: 120
Query Match: 6.8% Indels: 36
DB: 4 Gaps: 8

US-10-768-158-1 (1-2419) x US-10-092-900A-150 (1-295)
Qy 30 AGCGAGCGCGAGCCCGGAGGAGTTCGAGAGCAAGTACTTCGAGTTCAT 89
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 7 SerGluLeuGluTrpGlyLeuLeuProGlu-----GluPheSerGlnValAsn 23
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 90 GCGCTGCGGCTCGCCCGCTTCTGCGCGGGAAGATGGAGAGATCGCAACTTCCCGGTG 149
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 24 GlyIleLeuLeuGlnLysMetCysAspPheTrpAspLysIleTrpAsnPheGlnAla 43
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 150 CGGCCCGAGCGTGTGATCGTACCTACCCAGTCCGCGAGGAGTTCCTCGAGGAG 209
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 44 LysProAspAspLeuLeuAlaSerTyrProLysAlaGlyThrTrpThrGlnGlu 63
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 210 GTGCTACTTGGTGAGCGCGGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 64 IleValAspLeuIleGlnAsnAspGlyAspIleGluLysSerArgAlaSerIleGln 83
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 270 GAGCAGCTCCCGTCTCGAGTACCCAG-----CCGGGC 305
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 84 LeuGlnHisProPheLeuGluTrpIleArgMetThrHisAlaArgLysIlePheAlaGly 103
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 306 CTGACATCATCAGGAACCTACCTCCCGCTCATCAGAGCCACCTGCCCTACCGC 365
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 104 IleAspGlnAlaAsnThrMetProSerProArgThrLeuLysThrHisLeuProValGln 123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 366 TTCTGCGCTCTGACCTCCACATGGAGCTCCAGGTCTATATGCTCGCAACCCC 425
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 124 LeuLeuProProSerPheTrpGluGlnAsnCysIleIleTyrValAlaArgAsnAla 143
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 426 AAGGATCTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGTACCGA 485
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 144 LysAspAsnLeuValSerTyrTyrHisPheGlnArgMetSerLysAlaLeuProAspVal 163
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 486 GGCACCTTTCAGAAATTCGCGGAGGTTTATGATGATAGCTGGGTACGCTCCG 545
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 164 LeuThrValGlyGluTyr-----IleMetCysGlyGluValLeuTrpGlyIleTrp 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 546 TTTGAGCAGCTGCAGGAGTCTTGGGAGCACCAGCTGAGTCTGCAACGCTCTTCTCAAG 605
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

US-11-097-143-11370
; Sequence 11370, Application US/11097143
; Publication No, US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11370
; LENGTH: 313
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-11370

Alignment Scores:
Pred. No.: 1.53e-15 Length: 313
Score: 295.00 Matches: 83
Percent Similarity: 48.7% Conservative: 51
Best Local Similarity: 30.2% Mismatches: 103
Query Match: 6.7% Indels: 38
DB: 6 Gaps: 8

US-10-768-158-1 (1-2419) x US-11-097-143-11370 (1-313)
Qy 132 ATGCCCAACTTCCCGTGGCGGCCCGAGCGTGTGGATCGTACCTACCCCAAGTCCGGC 191
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 38 ValHisAspMetLysLeuArgAspAspValTrpIleValThrLeuProLysCysGly 57
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

181 ---GluGluIleArgThrTrpGlnLeuHisArg-----LeuPheCysTrp 194
Qy 606 TATGAGACATGATCGGACCTGGTGCACATGGTGAGCAGCTGGCCAGATTCCTCGGG 665
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
195 PhePheAspHisAlaSerGluAsnProArgLysPheLysArgIleMetGluPheMetGly 214
Qy 666 GTGTCCTGTGACCAAGGCCAGCTGGAGCCCTGACGAGCAGCTGC-----CACAG 716
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
215 AsnLysLeuAspGluAspProValLysArgIleValGlnHisThrSerPheGluSerLys 234
Qy 717 CTGGTGGACCATGCTGCAACGCTGAGCCCTGCC-----752
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
235 LysLysAsnGlnMetThrAsnTyrValMetIleThrCysAspIleMetAspHisSerIle 254
Qy 753 -----GTGGCCCGGGAAGAGTGGCTGTGAAGACATCTTACCGTCTCCATG 803
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
255 SerProPheMetArgLysGlyThrValGlyGluTrpLysAspTyrPheSerAlaAlaGln 274
Qy 804 AATGAGAGTTTGCATTTGGTGTATAAACAGAAAGTGTGACCTCACGTTT 860
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
275 AsnLysArgPheAsp-----GluAspArgLysMetAlaAspSerSerLeuThrPhe 291

RESULT 81
US-11-097-143-11370
; Sequence 11370, Application US/11097143
; Publication No, US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11370
; LENGTH: 313
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-11370
```



```

Qy 780 AAGGACATCTCACCGTCTCCATGAATGAGAAGTTTGACTTGTGTATAAACAG 833
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 229 LysAsnLeuPheSerGluIleGlnAsnGlnGluMetAspGluLysPheLysGlu 246
US-10-364-774-6
; Sequence 6, Application US/10364774
; Publication No. US2003014497A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/10/364,774
; PRIOR FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 265
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774-6

Alignment Scores:
Pred. No.: 2,338-15 Length: 265
Score: 292.50 Matches: 70
Percent Similarity: 49.2% Conservative: 57
Best Local Similarity: 27.1% Mismatches: 90
Query Match: 6.6% Indels: 41
DB: 4 Gaps: 7

US-10-768-158-1 (1-2419) x US-10-364-774-6 (1-265)
Qy 126 GAGGATCGCAACTCCCGTGGCGCCGACGACGTGTGATCGTCACCTACCCCAAG 185
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 8 GlnAlaLeuAspThrPheGluAlaArgHisAspAspIleValLeuAlaSerTyrProLys 27
Qy 186 TCCGCGCACAGC-----TTGCTGCGAGAGGTGGTCTACTTGTGAGCAGCGGC 233
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 28 CysGlySerAsnTrpIleLeuHisIleValSerGluLeuIleTyrAlaValSerLys--- 46
Qy 234 GCTGACCCCGATGAGATCGGCTTGATGAACATCGACGACGCTCCCGTCTCGGAGTAC 293
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 47 -----LysLysTyrLysTyr 51
Qy 294 CCACAG---CCGGGCTGACATC-----ATCAAGGAACGTG 326
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 52 ProGluPheProValLeuGluCysGlyAspSerGluLysTyrGlnArgMetLysGlyPhe 71
Qy 327 ACCTCTCCCGGCTCATCAAGACGACCTGCGCTTCTGCTGCTGACCTCCAC 386
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 72 ProSerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePhe 91
Qy 387 AATGAGAGACTCCAAAGTTCATCTATATGCTGCTCGCAAGGACTCTGTGTGTCTTAT 446

```

```

Db 92 GluAsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPhe 111
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 447 TATCAGTTCCACCGCTCTCTCGCGACCATGAGCTACCGAGCACCTTTCAAGATTTCGC 506
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 112 LeuHisPheHisAsnAspValProAspIleProSerTyrGlySerTyrAspGluPhePhe 131
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 507 CGGAGGTTTATGATGATAAGCTGGGCTACGGCTCCTGGTTGAGCAGCTGCAGGACTTC 566
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 132 ArgGlnPheMetLysGlyGlnValSerTrpGlyArgTyrPheAspPheAlaIleAsnTrp 151
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 567 TGGGAGCAGCCGATGAGCTCGAACGCTGTTTCTCAAGTATGAAGCATGCATCGGGAC 626
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 152 AsnLysHisLeuAspGlyAspAsnValLysPheIleLeuTyrGluAspLeuLysGluAsn 171
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 627 CTGCTGACGATGTTGGAGCAGCTGGCCAGATTCTCTGGGGGTCTCTCTGTGACAAAGGCCAG 686
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 172 LeuAlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPhePheLeuThrGlyGluGln 191
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 687 CTGGAAGCCCTGACGGAGCAGCTGCCACAGCTGGTGGACAGCTGCTCAACGCTGAGGCC 746
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 192 IleGlnThrIleSerVal-----GlnSerThrPheGlnAlaMetArgAlaLysSer 208
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 747 CTG-----CCCGTGGGC-----CGGGAGAGAGTTGGGCTGTGG 779
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 209 GlnAspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyGluValGlyAspTrp 228
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 780 AAGGACATCTTCACCGCTCCATGAATGAGAAGTTTGACTTGTGTATAAACAG 833
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 229 LysAsnLeuPheSerGluIleGlnAsnGlnGluMetAspGluLysPheLysGlu 246
RESULT 84
US-11-134-241-6
; Sequence 6, Application US/11134241
; Publication No. US20050287568A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/11/134,241
; PRIOR FILING DATE: 2005-05-20
; PRIOR APPLICATION NUMBER: US/10/364,774
; PRIOR FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 265
; TYPE: PRT
; ORGANISM: homo sapiens
US-11-134-241-6

Alignment Scores:
Pred. No.: 2,338-15 Length: 265

```



Score: 292.50 Matches: 70  
Percent Similarity: 49.2% Conservative: 57  
Best Local Similarity: 27.1% Mismatches: 90  
Query Match: 6.6% Indels: 41  
DB: 6 Gaps: 7

US-10-768-158-1 (1-2419) x US-11-134-241-6 (1-265)

Qy 126 GAGGAGATCGCAACTTCCGGTGGCGCCAGCGAGTGTGGATCGTCACTACCCCAAG 185  
Db 8 GlnAlaLeuAspThrPheGluAlaArgHisAspAlleValLeuAlaSerTyProlys 27  
Qy 186 TCCGCCACCGC-----TTGCTGAGGAGGTGGTCTACTTGGTGAGCCAGGC 233  
Db 28 CysGlySerAsnTrpIleLeuHisIleValSerGluLeuIleTyAlaValSerLys--- 46  
Qy 234 GCTGACCCCGATGAGATCGGCTTGATGAACATCGAGCAGCTCCCGTCTCTGGAGTAC 293  
Db 47 -----LysLysTyrlsTyrls 51  
Qy 294 CCACAG---CCGGGCTGGACATC-----ATCAAGGAATCG 326  
Db 52 ProGluPheProValLeuGluCysGlyAspSerGluLysTyrlsGlnArgMetLysGlyPhe 71  
Qy 327 ACCTTCCCGCTCATCAAGAGCACCTGCCCTTCTGCTTCTGACCTCCAC 386  
Db 72 ProSerProArgIleLeuAlaThrHisLeuHisTyrlsAspLysLeuProGlySerIlePhe 91  
Qy 387 AATGAGACTCCAAGTCTATATGCTCGCAACCCCAAGGATCTGGTGTCTTAT 446  
Db 92 GluAsnLysAlaLysIleValIlePheArgAsnProLysAspThrAlaValSerPhe 111  
Qy 447 TATCAGTTCCACCGCTCTCTGGGACCATGAGTACCGAGGACCTTTCACAAATTTGTC 506  
Db 112 LeuHisPheHisAsnAspValProAspIleProSerTyrlsGlySerTrpAspGluPhePhe 131  
Qy 507 CGGAGTTTATGATGATAAGCTGGCTACCGCTCTGCTGTTGAGCAGCTGCAGAGTTC 566  
Db 132 ArgGlnPheMetLysGlyGlnValSerTrpGlyArgTyrlsPheAspPheAlaIleAsnTrp 151  
Qy 567 TGGGAGCACCAGTGCAGTCCGACGTCTTTCTCAAGTATGAGACATGCATCGGC 626  
Db 152 AsnLysHisLeuAspGlyAspAsnValLysPheIleLeuTyrlsGluAspLeuLysGluAsn 171  
Qy 627 CTGTGTGACGATGCTGGAGCAGTGCAGATTCCTGGGGGTCTCTGTGACAAAGCCCGAC 686  
Db 172 LeuAlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPhePheLeuThrGlyGluGln 191  
Qy 687 CTGAGCCCTGACGAGGACACTGCACACGAGTGTGGACCGAGTGTGCAACGCTGAGGCC 746  
Db 192 IleGlnThrIleSerVal-----GlnSerThrPheGlnAlaMetArgAlaLysSer 208  
Qy 747 CTG-----CCCGTGGC-----CGGGAGAGTGGGTGG 779  
Db 209 GlnAspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyGluValGlyAspTrp 228  
Qy 780 AAGGACATCTTCACCGTCTCCATGATGAGAGTTTGTGTTGATATAAACAG 833  
Db 229 LysAsnLeuPheSerGluIleGlnAsnGlnGluMetAspGluLysPheLysGlu 246

RESULT 85  
US-10-072-012-665  
; Sequence 665, Application US/10072012  
; Publication No. US20040033493A1  
; GENERAL INFORMATION:  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shinkets, Richard  
; APPLICANT: Li, Li  
; APPLICANT: Gangolli, Esha  
; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Anderson, David W.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Taupier Jr, Raymond J.  
; APPLICANT: Gusev, Vladimir Y.  
; APPLICANT: Colman, Steven D.  
; APPLICANT: Wolenc, Adam R.  
; APPLICANT: Pena, Carol E. A  
; APPLICANT: Furtak, Katarzyna  
; APPLICANT: Grosse, William M.  
; APPLICANT: Alsbrook II, John P.  
; APPLICANT: Lepley, Denise M.  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Burgess, Catherine E.  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-258  
; CURRENT APPLICATION NUMBER: US/10/072,012  
; CURRENT FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: 60/265,102  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 60/265,514  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,517  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,412  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,395  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/266,406  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 60/266,767  
; PRIOR FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: 60/267,057  
; PRIOR FILING DATE: 2001-02-07  
; PRIOR APPLICATION NUMBER: 60/266,975  
; PRIOR FILING DATE: 2001-02-07  
; PRIOR APPLICATION NUMBER: 60/267,459  
; PRIOR FILING DATE: 2001-02-08  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1391  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 665  
; LENGTH: 312  
; TYPE: PRT  
; ORGANISM: Gallus gallus  
US-10-072-012-665

Alignment Scores:  
Pred. No.: 4.43e-15 Length: 312  
Score: 289.50 Matches: 69  
Percent Similarity: 50.4% Conservative: 68  
Best Local Similarity: 25.4% Mismatches: 122  
Query Match: 6.6% Indels: 13  
DB: 4 Gaps: 6

US-10-768-158-1 (1-2419) x US-10-072-012-665 (1-312)

Qy 78 TTCAGTTCATCGCGTGGCGTCCG---CCCTTCTGCGCGGGAAGATG---GAGGAG 131  
Db 29 PheSerTyrlsGlyValLeuTyrlsProValAlaLeuCysSerProGluValPheArgAla 48  
Qy 132 ATCGCCCAACTTCCCGTGGCGCCAGCGAGTGTGGATCGTCACTACCCCAAGTCCGGC 191  
Db 49 MetGluSerPheGluAlaArgSerAspValIleLeuAlaGlyTyrlsProLysSerGly 68  
Qy 192 ACCAGTGTGTGAGGAGGTGGTCTACTTGTGAGCAGCGCGCTGACCCCGATGAGATC 251  
Db 69 ThrAsnTrpValGlyGlnIleLeuSerAspLeuValAlaThrPheGluLysGluArgLeu 88  
Qy 252 GGCTTGATGAACATC---GACGAGCAGCTCCCGTCTCTGGAGTACCCACAGCCCGGC--- 305  
Db 89 GluLysSerValAsnAspGluGluPheProTyrlsLeuGluIleGlyAsp 108





```
OTHER INFORMATION: Ceres Seq. ID no. 15218612
US-11-096-568A-1016

Alignment Scores:
Pred. No.: 5,18e-14 Length: 345
Score: 277.00 Matches: 84
Percent Similarity: 42.8% Conservative: 43
Best Local Similarity: 28.3% Mismatches: 127
Query Match: 6.3% Indels: 43
DB: 6 Gaps: 10

US-10-768-158-1 (1-2419) x US-11-096-568A-1016 (1-345)
QY 53 CCGCGGGAGTTCGAGACAGTACTTCGAGTTCCA-----TGGCGTGGCGTCCGCC 106
Db 39 ProSerGlyThrLeuGlnAspTyrCysArgHisProSerGlyTrpCysIleThrLeu-Pr 58
QY 107 CTTTCGCCGGGAAGATGGAGGAGATCGCCAACTTCCCGTGGCGCCAGCAGCGTGTG 166
Db 58 oileMetValSerSerMetValAlaGluGlnHisPheGluAlaArgGlyThrAspValLe 78
QY 167 GATCGTCACTACCCCAAGTCGGCACCGAGTTCGTGCGAGGAGTGTCTACTTGGTGAG 226
Db 78 uLeuValThrMetProLysSerGlyThrThrTriPileLysAlaLeuLeuTyAlaAla 98
QY 227 CCAGGGCGCTGACCCCGATGAGATCGGCTTGATG-----AACATCGA 268
Db 98 aHisArgThrAspAspThrSerSerSerIleLeuArgGlnLeuAlaSerHisAsnSerHi 118
QY 269 CGAGCAGCTCCGGTCTCGGAG-----TACCACAGCGGGCCCTGGACATCATCAA 319
Db 118 sGlnLeuValProPheLeuGluAlaGlnValTyrThrLysAspGlnIleProAspLeuSe 138
QY 320 GAAACTGACCTTCCCGCCTCATCAAGAGCCACCTGCGCTTACCGCTTTCGCCCTCTGA 379
Db 138 rSerLeuProAlaProArgLeuPheAlaThrHisIleProAlaGluSerLeuProProSe 158
QY 380 CTTCCACATGGAGACTCCAGGTCATCTATATGCTCGCAACCCAGGATCTGGTGGT 439
Db 158 rValValAlaSerGlyCysLysValValTyrLeuCysArgAspProLysAspCysPheVa 178
QY 440 GTCTATTATCATGTTCCACCGCTCTCGGACCATGAGCTACCGAGGACCTTTCAAGA 499
Db 178 lSerLeuTrpHisPheMetAsnLysPheThrProTrpAspIleAspGluAlaHisGlyAr 198
QY 500 ATTTCGCGGAGTTTATGAATGATGAAGCTGGGC-----TACGGCTCTCGTTTGA 550
Db 198 gPheCysGlu-----GlyValSerLeuTyrGlyProPheTrpGl 211
QY 551 GCACGTGAGAGTCTGGGACCGCATGAGC-----TCGAACGTCTTTTCTCAA 604
Db 211 uHisValLeuSerTyrTrpArgTrpHisValAspArgProGlyGlnValLeuPheLeuTh 231
QY 605 GTATCAAGACATGATCGGACCTCGTGACGATGTGGAGCAGCTGGCCAGATTCTCTGG 664
Db 231 rTyrGluGluLeuSerAlaAspProLeuGlyGlnLeuArgLeuAlaGluPheIleGl 251
QY 665 GGTGTCCTGTGAAGGCCAGCTCGAAGCC-----CTGACGGACCATG 709
Db 251 yArgProPheThrProGlyGluGlnGluAlaGlyValAspArgGluIleAlaGluAlaCy 271
QY 710 C-----CACCAGCTGGTGACCGAGTCTGCTCAACGCTGAG----- 743
Db 271 sAlaMetLysSerMetValAsnGlnGluValAsnGlnSerArgThrThrGluIleValGl 291
QY 744 -----GCCCTCGCGTGGG-----CGGGAGAGAGTTGGGCTGTGGAAGACAT 787
Db 291 uMetProIleProAsnGlyLysPhePheArgArgGlyValValGlyAspTrpThrAsnTy 311
QY 788 CTTCCCGCTCCATGAATGAGAAGTTGACTTGGTGTGTATTAACAGAG 836
Db 311 rLeuThrProGluMetAlaGlyArgIleAspGluIleThrLysSerLys 327
```

```
RESULT 89
US-09-854-122-20
; Sequence 20 Application US/09854122
; Patent No. US20020016980A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTE, RANDALL S.
; APPLICANT: SMITH, ROBERT
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
; FILE REFERENCE: PHA-007.01
; CURRENT APPLICATION NUMBER: US/09/854,122
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/202,529
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; TYPE: PRT
; ORGANISM: Flaveria bidentis
US-09-854-122-20

Alignment Scores:
Pred. No.: 6,74e-14 Length: 320
Score: 275.50 Matches: 80
Percent Similarity: 49.1% Conservative: 63
Best Local Similarity: 27.5% Mismatches: 111
Query Match: 6.3% Indels: 37
DB: 3 Gaps: 11

US-10-768-158-1 (1-2419) x US-09-854-122-20 (1-320)
QY 72 AAGTACTTCGAGTTCATCGGCTGGCGTCCGCCCTTCTGCCGGAAGATGAGGAG 131
Db 34 LysTyrGlnAspPheTrpGly-----LeuGlnAsnIleGluGlyAlaIleLeuAla 51
QY 132 ATCCCAACTTCCCGTGGCGCCAGCAGCGTGTGATCGTCACTACCCCAAGTCCGGC 191
Db 52 GlnGlnSerPheLysAlaArgProAspValPheLeuCysSerTyrProLysSerGly 71
QY 192 ACCAGTGTGTGAGGAGTGTCTACTTGGTGGCAGCGGCGTGCACCCGATGAGATC 251
Db 72 ThrThrTrpLeuLysAlaLeuAlaTyrAlaIleValThrArgGluLysPheAspGluPhe 91
QY 252 GGC-----TTGATGAACATCAGCAGCAGCAG-----CTCCCGTCTGGAGTACCCACAG 299
Db 92 ThrSerProLeuLeuThrAsnIleProHisAsnCysIleProTyrIleGlu----- 108
QY 300 CCGGCGCTGGACATCAAGGAACTGACCTCTCCCGC-----CTCATC 344
Db 109 -----LysAspLeuLysLysIleValGluAsnGlnAsnSerCysPheThrProMet 126
QY 345 AAGAGCCACCTGCGCTACCGCTTTCGCCCTCTGACCTCCACATGGAGACTCCAGTGC 404
Db 127 AlaThrHisMetProTyrHisValLeuProLysSerIleLeuAlaLeuAsnCysLysMet 146
QY 405 ATCTATATGGCTCGCAACCCCAAGATCTGGTGTGTCTTATTATCATCTCCACCGCTCT 464
Db 147 ValTyrIleTyrArgAsnIleLysAspValIleValSerPheTyrHisPheGlyArgGlu 166
QY 465 CTGCGGACCATGAGCTACCGA---GGCACCTTTCAAGAAATTCGCGGAGGTATTGAAT 521
Db 167 IleThrLysLeuProLeuGluAspAlaProPheGluGluAlaPheAspGluPheTyrHis 186
QY 522 GATTAAGCTGGGCTACCGCTCTCTGGTTTGGACGCTGGAGAGTTCGGGACCGCATG 581
Db 187 GlyIleSerGlnPheGlyProTyrTrpAspHisLeuLeuGlyTyrTrpIleAspLeu 206
QY 582 GAC-----TCGAACGTGCTTTTCTCAAGTATGAAGACATGCATCGGAGCTGTGTGACG 635
Db 207 GluArgProGluValIleLeuPheLeuLysTyrGluAspValLysLysAspProThrSer 226
QY 636 ATGTGTGAGCAGCTGGCCAGATTCCTGGGGGTGTCTCTGT-----CACAG 680
```

Db      227 AsnValLysArgLeuAlaGluPheIleGlyTyrProPheThrPheGluGluLysGlu    24

Qy      681 GCCCAGCTCGGAAGCCCTCACGGAGCACTGC-----CACCAGCTGTGTG    722  
         :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db      247 GlyValIleGluSerIlellellysLeuCysSerPheGluAsnLeuSerAsnLeuVal    266  
         :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Qy      723 GACCACTGCTCCAACGCTGAGGCC---CTGCCCGTG------GCCCGG    761  
         :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db      267 AsnLysSerGlyAsnSerLysGlyPheLeuProIleGluAsnArgLeuTyrPheArgLys    286  
         :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Qy      762 GGAAGAGTTGGCTCTGGAAGGACATCTTCACCGTCCCATGAATGAGAAGTTTGCATTG    821  
         :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db      287 AlaLysAspGlyAspTrpLysAsnTyrPheThrAspGluMetThrGluLysIleAspLys    306  
         :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Qy      822 GTGTATAAACAGAAGATGGAAAGTGTGACTTC    854  
         :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db      307 LeuIleAspGluLysLeuSerAlaThrGlyLeu    317  
         :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
  
RESULT 90  
US-11-033-030-20  
; Sequence 20, Application US/11033030  
; Publication No. US20060053510A1  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTE, RANDALL S.  
; APPLICANT: SMITH, ROBERT  
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA  
; FILE REFERENCE: PHA-007.01  
; CURRENT APPLICATION NUMBER: US/11/033,030  
; CURRENT FILING DATE: 2005-01-11  
; PRIOR APPLICATION NUMBER: US/09/854,122  
; PRIOR FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: 60/202,529  
; PRIOR FILING DATE: 2000-05-10  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 320  
; TYPE: PR1  
; ORGANISM: Flaveria bidentis  
US-11-033-030-20  
  
Alignment Scores:  
Pred. No.:          6.74e-14          Length:          320  
Score:             275.50          Matches:         80  
Percent Similarity: 49.1%        Conservative:    63  
Best Local Similarity: 27.5%      Mismatches:    111  
Query Match:       6.3%           Indels:          37  
DB:                 6                Gaps:            11  
  
US-10-768-158-1 (1-2419) x US-11-033-030-20 (1-320)  
Qy      72 AAGTACTTCGATTCCATGGCGTGCGGTGCGGCCCTTCTGCGCGGGAAGATGAGGAG    131  
         ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db      34 LysTyrGlnAspPheTrpGly-----LeuGlnAsnAsnIleGluGlyAlaIleLeuAla    51  
         :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Qy      132 ATCCCAAATCTCCGGTGCGGCCCAACGACGACGTGTGGATCGTCACTACCCCCAAGTCCC GC    191  
         :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db      52 GlnGlnSerPheLysAlaArgProAspAspValPheLeuCysSerTyrProLysSerGly    71  
         :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Qy      192 ACCAGCTTGTGCAGGAGGTGTTCTACTTGGTAGCCAGGCGGTGACCCCGATGATGATC    251  
         ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db      72 ThrThrTrpLeuLysAlaLeuAlaTyrAlaIleValThrArgGluLysPheAspGluPhe    91  
         :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Qy      252 GGC-----TTGATCAACATCGACGAGCAG-- -CTCCCGGTCTCTGGAGTACCACAG    299  
         ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db      92 ThrSerProLeuLeuThraSnnIleProHisAsnCysIleProTyrlleGu -----    108  
         :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Qy      300 CCGGGCCTCGGACATCATCAAGGAATGACTACCTCTCCCCGC-----CTCATC    344  
         ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db      109 -----LysAspLeuLysLysIleValGluAsnGlnAsnAsnSerCysPheThrProMet    126  
         :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Qy      345 AAGAGCCACTTGCCCTACCGCTTTCTGCCCCTCTGACCTCCACAATGGAGACTCCAAGTTC    404  
         :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```
Qy 39 GAGACCCACACCCCGGGGAGTTCGAGAGCAAGTACTTCGAGTTCCTCCATGGCGTGGC 98
Db 27 ---ValProAlaLeuProLeuGlu-----ThrArg 35
Qy 99 CTGCGCGCTCTCTGC----- 113
Db 36 CysProProPheAlaLeuArgArgTyxGlyGlyPheTrpLeuProGluAlaLeuArg 55
Qy 114 CGCGGGAAGATGGAGGAGATGCCCACTTCGCGTGGCGCCAGCAGCTGTGGATGTC 173
Db 56 ArgGlyLeuProAlaLeuHisAlaGlyPheAlaProArgArgAlaAspValLeuLeuAla 75
Qy 174 ACCTACCCCAAGTCGCGCACACAGCTTGTGAGAGGTGTCTACTTGGTGGAGCCAGGC 233
Db 76 SerTyProLysSerGlyThrTrpLeuLysAlaLeuPheAlaThrValArgArg 95
Qy 234 GCT-----GACCCCGATGAGATCGGC-----TTGATGAACATCGACGAGCAG 275
Db 96 AlaValHisProProSerSerIleAspHisProLeuArgArgArgProGlnAspCys 115
Qy 276 CTCCCGTCTCGAGTACCCACAGCCCGGCTGGACATC-----ATCAAGGAACCTGACC 329
Db 116 ValProPheLeuGluPheAspPheAspHisGlyAspMetAlaGluLeuGluAlaLeuPro 135
Qy 330 TCTCCCGCTCATCAGAGCACCTGCCCTACCGCTTCTGCGCTCTGACCTCCACAAT 389
Db 136 SerProArgValLeuAlaThrHisLeuProTyProSerLeuLeuProAspArgValVala 155
Qy 390 GGAGAC---TCCAAAGTCACTATATGCTCGCAACCCCAAGGATCTGGTGTCTTAT 446
Db 156 AspAspGlyAlaArgIleValTyValCysArgAspProLysAspAlaLeuValSerSer 175
Qy 447 TATCAGTTCCACCG-----TCTCTCGGACCATGAGCTACCGAGGACCTTT 494
Db 176 TrpLeuPheThrArgLysAlaSerProSerValGlyValAspSerAspSerTySerLeu 195
Qy 495 CAAGAATTCCTCCGAGGTTTATGAATGATGAAGCTGGCTACCGCTCTGTTGGAGCAC 554
Db 196 GlnGluAlaAlaGluLeuPheArgGluGlyArgCysPheTyProGlnCysArgHis 215
Qy 555 GTGAGGAGTTCGGGACACCGCATG-----GACTCGAAGTCTCTTTTCTC 602
Db 216 ValLeuGluTyTrpAspAlaSerArgArgProPheGlyGlyGlyValLeuPheLeu 235
Qy 603 AAGTATGAAGACATCGCATCGGACCTGTGTGACGATGTGTGGAGCAGTGGCCAGATTCTG 662
Db 236 ArgTyProGluMetLeuArgAspProAlaGlySerLeuLysThrMetAlaGluPheMet 255
Qy 663 GGGGTGCTGTGACAGGCCACGCTGGAAGCCCTGACGGAGCAGCTGCCACCACTGGTG 722
Db 256 Gly-----CysGlyPheSerGluGluGluValAlaArgGlyValValAspGluIleVal 273
Qy 723 GACCACTGCTGC-----AACGCTGAGGCCCTGCCGTCGGCGCG----- 761
Db 274 AspLeuCysSerLeuGluLysLeuLysAsnMetGluAlaAsnArgAspGlyArgArgAsn 293
Qy 762 -----GGAAGATTTGGCTGTGGAAGGAC 785
Db 294 AlaSerGlyIleArgSerAspSerPhePheArgLysGlyValAlaGlyAspTrpSerSer 313
Qy 786 ATCTTCACCGTCTCCATGAATGAGAGTTTACATTGGTGTATAAACAGAAATGGAAAG 845
Db 314 HisMetSerProGluMetGlyLysMetLeuAspGluAlaValGluGluAlaLeuGlnGly 333
Qy 846 TGTGACCTCACGTTT 860
Db 334 SerGlyPheSerPhe 338
```

RESULT 92

US-11-096-568A-1018

; Sequence 1018, Application US/11096568A

; Publication No. US20060048240A1

## ; GENERAL INFORMATION:

```
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 1018
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(282)
; OTHER INFORMATION: Ceres Seq. ID no. 15218614
US-11-096-568A-1018
```

## Alignment Scores:

```
Pred. No.: 8,59e-14 Length: 282
Score: 274.00 Matches: 77
Percent Similarity: 44.7% Conservative: 42
Best Local Similarity: 28.9% Mismatches: 107
Query Match: 6.2% Indels: 40
DB: Gaps: 9
```

US-10-768-158-1 (1-2419) x US-11-096-568A-1018 (1-282)

```
Qy 138 AACTTCCCGTGGCGCCAGCAGCTGTGGATCGTCACTTACCCCAAGTCCGGCACCAGC 197
Db 6 HisPheGluAlaArgGlyThrAspValLeuLeuValThrMetProLysSerGlyThrThr 25
Qy 198 TTGCTGAGAGGTGTCTACTTGTGGTGGCAGCGGCGCTACCCCGATGAGATCGGTTG 257
Db 26 TrpIleLysAlaLeuLeuTyAlaAlaHisArgThrAspThrSerSerSerIle 45
Qy 258 ATG-----AACATCGACGAGCAGCTCCCGTCTCGAG----- 290
Db 46 LeuArgGlnLeuAlaSerHisAsnSerHisGlnLeuValProPheLeuGluAlaGlnVal 65
Qy 291 TACCACAGCGCGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGC 350
Db 66 TyrThrLysAspGlnIleProAspLeuSerSerLeuProAlaProArgLeuPheAlaThr 85
Qy 351 CACTGCGCTACCGCTTCTGCGCTCTGACCTCCCAATCGAGACTCCAAGTCTCATAT 410
Db 86 HisIleProAlaGluSerLeuProProSerValValAlaSerGlyCysLysValValTy 105
Qy 411 ATGCTCGCAACCCCAAGGATCTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGCG 470
Db 106 LeuCysArgAspProLysAspCysPheValSerLeuTrpHisPheMetAsnLysPheThr 125
Qy 471 ACCATGAGTACCGAGGCACCTTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTG 530
Db 126 ProTrpAspIleAspGluAlaHisGlyArgPheCysGlu----- 138
Qy 531 GGC-----TACGCGCTCTGTTGACACGCTGAGGAGTCTCGGACACCGCATG 581
Db 139 GlyValSerLeuTyArgProPheTrpGluHisValLeuSerTyTrpArgTrpHisVal 158
Qy 582 GAC-----TCGAACGTGCTTTTCTCAAGTATGAAGACATCATCGGAGCTCTGTCAGC 635
Db 159 AspArgProGlnValLeuPheLeuThrTyArgGluLeuSerAlaAspProLeuGly 178
Qy 636 ATGCTGAGCAGCTGGCCAGATCTCTGGGGGTGCTGTGACAAAGGCCAGCTCGAAGCC 695
Db 179 GlnLeuArgArgLeuAlaGluPheIleGlyArgProPheThrProGlyGluGlnGluAla 198
Qy 696 -----CTGACGGAGCACTGC-----CACCACCTGCTG 722
Db 199 GlyValAspArgGluIleAlaGluAlaCysAlaMetLysSerMetValAsnGlnGluVal 218
Qy 723 GACCAGTGTCTGCAACGCTGAG-----GCCCTGCCCGTGGCG----- 758
```



Db 84 -----AspGluHisHisProLeuLeuGluHisAsnProHisAspLeuValGly 99  
 Qy 297 -----CAGCGGGCTGGACATATCAAGAACTGACCTCTCCCGGCTCATC 344  
 Db 100 PheLeuGluIleTyrProLysLeuGluLeuTyrGluSerLeuProSerProArgLeuLeu 119  
 Qy 345 AAGAGCCACCTGCGCTTCTGCGCTCTGACCTCCACATGGAGACTCC----- 398  
 Db 120 SerThrHisLeuProTyrSerMetLeuProHisArgIleArgGluGlnGluThrGlyCys 139  
 Qy 399 AAGGTCTATATATGGTTCGCAACCCCAAGGATCTGGTGTCTTATATCACTTCCAC 458  
 Db 140 ArgLeuValTyrIleTyrArgAspProLysAspAlaMetValSer-----MetTriPHis 157  
 Qy 459 CGCTCTTCGGGACCATGAGTACCGAGGCACTTTCAAGAAATTCGCGGAGGTTTATG 518  
 Db 158 GlnAsnLysLysGluLysLysAsnArgLeuThrPheGluGluMetPheAspMetPheCys 177  
 Qy 519 AATGATAAGCTGGGCTACGGCTCTGTTTTCGACGCGTGCAGGAGTTCTGGGACGACCGC 578  
 Db 178 GluGlyArgCysValValGlyProGlnTrpCysHisAlaGlyGluTyrTrpAspLysSer 197  
 Qy 579 ATG-----GACTCGAAGCTCTTTTCTCAAGTATGAAGACATCGCATCGGACCTGTG 632  
 Db 198 GlnAlaArgProGluLysValLeuPheLeuMetTyrGluAspLeuLeuGlnAspThrVal 217  
 Qy 633 ACGATGTGGAGCAGCTGGCGAGATTCCTGGGGGTCTCTGTGACAGGCCACCGCTGAA 692  
 Db 218 GlyAsnLeuArgThrLeuAlaGluPheMetGly-----CysGlyPheSerArgGlnGlu 235  
 Qy 693 GCCCTGCGGAGACTGCCACGACTGGTGACGAGTGCACAGTCTGC-----AACGCTGAGGCC 746  
 Db 236 GluAspAspGlyIleValGlnGlnIleValGluLeuCysSerLeuAsnAsnLeuLysAsn 255  
 Qy 747 CTGCGCGTGGG----- 758  
 Db 256 LeuAsnValAsnLysSerGlyThrThrLeuLeuGlyIleSerLysAspGlyPheArg 275  
 Qy 759 CGGGAGAGAGTTGGCTGTGAAGACATCTTCACCGTCTCCATGAATGAGAAGTTTGAC 818  
 Db 276 LysGlyGlyThrGlyAspTrpSerAsnHisMetSerProGluMetAlaAlaArgLeuAsp 295  
 Qy 819 TTGGTGTTATAACAGAAATG 839  
 Db 296 LysIleValLysGluArgLeu 302

## RESULT 95

US-11-096-568A-7712  
 ; Sequence 7712, Application US/11096568A  
 ; Publication No. US20060048240A1

## GENERAL INFORMATION:

; APPLICANT: Alexandrov, Nikolai et al.  
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
 ; FILE REFERENCE: 2750-1592PUS2  
 ; CURRENT APPLICATION NUMBER: US/11/096,568A  
 ; CURRENT FILING DATE: 2005-04-01  
 ; NUMBER OF SEQ ID NOS: 34471

SEQ ID NO 7712

LENGTH: 329

TYPE: PPT

ORGANISM: Glycine max

FEATURE:

NAME/KEY: misc.feature

LOCATION: (1)..(329)

OTHER INFORMATION: Ceres Seq. ID no. 15174866

US-11-096-568A-7712

## Alignment Scores:

Pred. No.: 2,64e-13 Length: 329  
 Score: 268.50 Matches: 78  
 Percent Similarity: 44.2% Conservative: 45  
 Best Local Similarity: 28.1% Mismatches: 106

Query Match: 6.1% Indels: 49  
 DB: Gaps: 9

US-10-768-158-1 (1-2419) x US-11-096-568A-7712 (1-329)

Qy 141 TTCCCGGTGGCGCCAGCGACGTGTGATCGTCACCTACCCCAAGTCCGCCAGCAGCTTG 200  
 Db 63 PheValProArgAlaAspAspIleIleLeuAlaThrGlnProLysCysGlyThrThrTrp 82  
 Qy 201 CTCGAGGAGTGTCTACTTGGTGAGCCAG-----GGCGGTGACCCCGATGAG 248  
 Db 83 LeuLysAlaLeuAlaPheThrIleAlaAsnArgSerArgTyrGlyPheSerAspHisPro 102  
 Qy 249 ATCGGCTGTATGAACATCGACGACGCTCCGGTCTCGGAGTACCCA-----CAGCCG 302  
 Db 103 LeuLeuThrArgHisProGlnHisValValProPheIleGluIleProGlyAlaGlyPro 122  
 Qy 303 GGCCTGACATCATCAAGGAACGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTAC 362  
 Db 123 AsnHisThrAspIleHisThrLeuProProArgLeuLeuAlaThrHisMetProMet 142  
 Qy 363 CGTTTCTGCCCTCTGACCTCCCAATGGAGACTCCAAAGTCACTTATATGTCGTCGCAAC 422  
 Db 143 SerLeuLeuProProGlyThrArgSerValGlyCysArgValValTyrLeuCysArgAsp 162  
 Qy 423 CCCAAGGATCTGGTGTCTCTTATTATCAGTTCCACCGCTCTCTCGCGGACCATGAGCTAC 482  
 Db 163 ProLysAspAlaLeuValSerArgLeuHisPheGluAsnLys-----AlaPhe 178  
 Qy 483 CGAGGCACC-----TTTCAAGAAATTCGCCGAGGTTTATGAAT 521  
 Db 179 GlnGlyThrAsnLeuSerMetAspAsnAlaPheSerMetPheCysGluGlyPheSer--- 197  
 Qy 522 GATAAGCTGGCTACGCTCTCTGTTGAGCATGTCGAGAGTTCTCGGAGCACCGCATG 581  
 Db 198 -----ProTyrGlyProPheTrpAspHisCysLeuGlyTyrTrpArgGluSerIle 214  
 Qy 582 -----CACTCGACGCTGCTTTTCTCAAGTATGAACATGATCGCGGACCTGGTCACG 635  
 Db 215 AlaArgProAspAsnValLeuPheLeuLysTyrGluIleLysSerAspProValPro 234  
 Qy 636 ATGTGTGAGCAGCTGGCGCAGATTCCTGGGGGTGCTCTGTCACAGGCCACCGCTGGAAGCC 695  
 Db 235 ValValArgLysLeuAlaLysPheLeuGlyValProLeuThrGlu-----GluGlu 251  
 Qy 696 CTGACGAGGACCTGCCACGCTGGTGGACATGTCGACAGCTGCAACGCTGAGGCCCTG 749  
 Db 252 GluSerSerGlyValAlaGlnGluValValLysLeuLysSerPheGluThrLeuThrSer 271  
 Qy 750 -----CCCGTGGGC----- 758  
 Db 272 LeuGlnValAsnGlnValGlyValArgHisGlyAspLysIlePheMetSerAsnSer 291  
 Qy 759 -----CGGGGAAGAGTTGGCTGTGGAAGGACATCTTCACCGTCTCCATGAAT 806  
 Db 292 ValPheTyrArgLysGlyGluValGlyAspTrpAlaAsnHisMetSerHisValMetGly 311  
 Qy 807 GAGAAGTTTGACTTGGTGTATAAACAGAAAGTGTGACCTACCTTTCACCTTT 860  
 Db 312 GluGluLeuAspArgIleValGlnGlnLysLeuGlySerGlyLeuValPhe 329

## RESULT 96

US-11-096-568A-7711  
 ; Sequence 7711, Application US/11096568A  
 ; Publication No. US20060048240A1

## GENERAL INFORMATION:

; APPLICANT: Alexandrov, Nikolai et al.  
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
 ; FILE REFERENCE: 2750-1592PUS2  
 ; CURRENT APPLICATION NUMBER: US/11/096,568A  
 ; CURRENT FILING DATE: 2005-04-01  
 ; NUMBER OF SEQ ID NOS: 34471



```
; SEQ ID NO 7711
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(330)
; OTHER INFORMATION: Ceres Seq. ID no. 15174865
US-11-096-568A-7711

Alignment Scores:
Pred. No.: 2.65e-13 Length: 330
Score: 268.50 Matches: 78
Percent Similarity: 44.2% Conservative: 45
Best Local Similarity: 28.1% Mismatches: 106
Query Match: 6.1% Indels: 49
DB: Gaps: 9

US-10-768-158-1 (1-2419) x US-11-096-568A-7711 (1-330)
Qy 141 TTCCGGTGGCCGCGCCAGCGTGTGGATCGTACCTACCCCAAGTCGGGACCGAGCTTG 200
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
64 PheValProArgAlaAspAspIleLeuAlaThrGlnProLysCysGlyThrTrp 83
Qy 201 CTGCAGGAGGTGCTACTTCTGTCAGCCAG-----GGCGCTGACCCCGATGAG 248
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
84 LeuLysAlaLeuAlaPheThrIleAlaAsnArgSerArgTyrGlyPheSerAspHisPro 103
Qy 249 ATCCGCTTGTATGAACATCGACGAGCTCCCGCTCTGGAGTACCCA-----CAGCG 302
Db :::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
104 LeuLeuThrArgHisProGlnHisValProPheIleGluIleProGlyAlaGlyPro 123
Qy 303 GGCTTGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTAC 362
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
124 AsnHisThrAspIleHisThrLeuProProArgLeuLeuAlaThrHisMetProMet 143
Qy 363 CGCTTTCTGCCTCTGACCTCCACATGGAGACTCCCAAGGTCACTATATGGCTCGCAAC 422
Db :::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
144 SerLeuLeuProGlyThrArgSerValGlyCysArgValValTyrLeuCysArgAsp 163
Qy 423 CCCAAGGATCGTGGTGTCTATTATCAGTTTCACCGCTCTCTGCGGACCATGAGTAC 482
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
164 ProlsAspAlaLeuValSerArgLeuHisPheGluAsnLys-----AlaPhe 179
Qy 483 CGAGGCACC-----TTTCAAGAATTCGCGGAGGTTTATGAAT 521
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
180 GlnGlyThrAsnLeuSerMetAspAsnAlaPheSerMetPheCysGluGlyPheSer--- 198
Qy 522 GATAAGCTGGCTACGCTCTCTGTTGAGCAGCTGCGAGGTTCTGGGACACCGCATG 581
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
199 -----ProTyrGlyProPheTrpAspHisCysLeuGlyTyrTrpArgGluSerIle 215
Qy 582 -----GACTCGAAGCTGTTTCTCAAGTATGAAGATGCATCGGACCTGGTGACG 635
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
216 AlaArgProAspAsnValLeuPheLeuLysTyrGluGluIleLysSerAspProValPro 235
Qy 636 ATGTGTGAGCAGCTGGCCAGATTCCTGGGGGTGTCCTGTGCAAGGCGCCAGCTGGAAGCC 695
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
236 ValValArgLysLeuAlaLysPheLeuGlyValProLeuThrGlu-----GluGlu 252
Qy 696 CTGACGGAGCACTGCCACCACTGGTGACCAAGTCTGCAACGCTGAGGCGCTG----- 749
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
253 GluSerSerGlyValAlaGlnGluValValLysLeuCysSerPheGluThrLeuThrSer 272
Qy 750 -----CCGTGGGC----- 758
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
273 LeuGlnValAsnGlnValGlyGlyValArgHisGlyAspLysIlePheMetSerAsnSer 292
Qy 759 -----CGGGGAAGAGTTGGGCTGTGGGAAGACATCTTCACCGTCTCCATGAAT 806
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
293 ValPheTyrArgLysGlyGluValGlyAspTrpAlaAsnHisMetSerHisValMetGly 312
Qy 807 GAGAAGTTTGACTTGGTGTATAAACAAGATGGAAAGTGTCGACCTTCACGTTT 860
```

```
Db 313 GluGluLeuAspArgIleValGlnGlnLysLeuGluClySerGlyLeuValPhe 330
RESULT 97
US-11-096-568A-7710
; Sequence 7710, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096.568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 7710
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(340)
; OTHER INFORMATION: Ceres Seq. ID no. 15174864
US-11-096-568A-7710

Alignment Scores:
Pred. No.: 2.68e-13 Length: 340
Score: 268.50 Matches: 78
Percent Similarity: 44.2% Conservative: 45
Best Local Similarity: 28.1% Mismatches: 106
Query Match: 6.1% Indels: 49
DB: Gaps: 9

US-10-768-158-1 (1-2419) x US-11-096-568A-7710 (1-340)
Qy 141 TTCCGGTGGCCGCGCCAGCGTGTGGATCGTACCTACCCCAAGTCGGGACCGAGCTTG 200
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
74 PheValProArgAlaAspAspIleLeuAlaThrGlnProLysCysGlyThrTrp 93
Qy 201 CTGCAGGAGGTGCTACTTCTGTCAGCCAG-----GGCGCTGACCCCGATGAG 248
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
94 LeuLysAlaLeuAlaPheThrIleAlaAsnArgSerArgTyrGlyPheSerAspHisPro 113
Qy 249 ATCCGCTTGTATGAACATCGACGAGCTCCCGTCTCGGAGTACCCA-----CAGCG 302
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
114 LeuLeuThrArgHisProGlnHisValProPheIleGluIleProGlyAlaGlyPro 133
Qy 303 GGCTTGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTAC 362
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
134 AsnHisThrAspIleHisThrLeuProProArgLeuLeuAlaThrHisMetProMet 153
Qy 363 CGCTTTCTGCCTCTGACCTCCCAATGGAGACTCCCAAGGTCACTATATGCTCGCAAC 422
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
154 SerLeuLeuProProGlyThrArgSerValGlyCysArgValValTyrLeuCysArgAsp 173
Qy 423 CCCAAGGATCGTGGTGTCTTATTATCAGTTTCACCGCTCTCTGCGGACCATGAGTAC 482
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
174 ProlsAspAlaLeuValSerArgLeuHisPheGluAsnLys-----AlaPhe 189
Qy 483 CGAGGCACC-----TTTCAAGAATTCGCGGAGGTTTATGAAT 521
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
190 GlnGlyThrAsnLeuSerMetAspAsnAlaPheSerMetPheCysGluGlyPheSer--- 208
Qy 522 GATAAGCTGGCTACGCTCTCTGTTGAGCAGCTGCGGAGTTCCTGGGACACCGCATG 581
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
209 -----ProTyrGlyProPheTrpAspHisCysLeuGlyTyrTrpArgGluSerIle 225
Qy 582 -----GACTCGAAGCTGTTTCTCAAGTATGAAGATGCATCGGACCTGGTGACG 635
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
226 AlaArgProAspAsnValLeuPheLeuLysTyrGluGluIleLysSerAspProValPro 245
Qy 636 ATGTGTGAGCAGCTGGCCAGATTCCTGGGGGTGTCCTGTGCAAGAGCCCGGAGGCC 695
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 246 ValValArgLysLeuAlaLysPheLeuGlyValProLeuThrGlu-----GluGlu 262
Qy 696 CTGACGGAGACTCCACACAGCTGGTGACAGCTGCGCAACGCTGAGGCCCTG----- 749
Db 263 GluSerSerGlyValAlaGlnGluValValLysLeuCysSerPheGluThrLeuThrSer 282
Qy 750 -----CCCGTGGGC----- 758
Db 283 LeuGlnValAenGlnValGlyValArgHisGlyAspLysIlePheMetSerAsnSer 302
Qy 759 -----CGGGAGAGATTGGCTGTGGAGGACACATCTTCCACCGTCCCATGAAT 806
Db 303 ValPheTyrArgLysGlyGluValGlyAspTrpAlaAenHisMetSerHisValMetGly 322
Qy 807 GAGAACTTGAATGGTGTTATAACACAGAGATGGAAAGTGTGACCTCACGTTT 860
Db 323 GluGluLeuAspArgIleValGlnGlnLysLeuGluGlySerGlyLeuValPhe 340

RESULT 99
US-10-767-701-42402
; Sequence 42402, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 42402
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C44289_1.pep
US-10-767-701-42402

Alignment Scores:
Pred. No.: 3,26e-13 Length: 344
Score: 267.50 Matches: 78
Percent Similarity: 45.5% Conservative: 47
Best Local Similarity: 28.4% Mismatches: 109
Query Match: 6.1% Indels: 41
DB: 4 Gaps: 10

US-10-768-158-1 (1-2419) x US-10-767-701-42402 (1-344)
Qy 141 TTCGCGGTGGCGCCAGCAGCGTGTGGATCGTCACCTACCCCAAGTCCGGCAGCAGCTTG 200
Db 75 PheThrProArgProAspAspValIleLeuAlaSerTyrProLysCysGlyThrThrTrp 94
Qy 201 CTGACGAGAGTGGTCTACTTGTGTGAGC-----CAGGCGCTGACCCCGATGAGATCGGC 254
Db 95 LeuLysAlaLeuAlaPheAlaThrAlaAlaAlaArgThrAlaTyrProProSerAspAlaGly 114
Qy 255 -----TTGATGAACATCGACGAGCAGCTCCCGCTCGGTGAGTACCCACAGCCG 302
Db 115 HisProLeuArgArgMetAsnProHisAspCysIleProPheIleAspGluIlePheAla 134
Qy 303 GGCCTGGACATCATCAAG-----GAACCTGACCTTCCCGCTCATCAAGACCACTG 356
Db 135 GlyGlyGluAspAlaLysLeuAspAlaLeuProSerProArgLeuMetAsnThrHisLeu 154
Qy 357 CCCTACCGCTTCTGCGCTCTGACCTC-----CACAAATGGAGCTCCCAAGTTCATCTAT 410
Db 155 ProTyrAlaLeuProProIleAspValValThrAlaThrGlyGlyCysArgValAlaTyr 174
Qy 411 ATGGCTCGCAACCCCAAGATCTGGTGTGTCTTATTATCATGTTCC---CACCGTCTCTG 467
Db 175 ValCysArgAspProLysAspMetValValSerLeuTrpHisPheLeuArgArgThrLys 194
```

```
Qy 468 CGACCATGAGCTACCGAGGCACTTTTCAAGAAATTCCTCGCGAGGTTTATGAATGATAAG 527
Db 195 ProAspLeuSerPheAlaAspThrPheGluSerValCys-----AspGlyThr 210
Qy 528 CTGGGCTACCGCTCTCTGTTTGGACAGCTGCAGAGTTCTGGGAGCACCGCATG----- 581
Db 211 ValAlaAlaGlyProValTrpAspHisValLeuSerTyrTrpArgAlaSerValAlaThr 230
Qy 582 GACTCGAACCTGCTTTTCTCAAGTATGAAGACATGATCGGACCTGGTGACCATGTTG 641
Db 231 ProAspArgAlaLeuPheLeuLysTyrGluAspLeuLeuArgAspProAlaGlyAsnVal 250
Qy 642 GAGCAGCTGCCAGATTCTCTGGGGGTCTCTGTGACAAAGCCCGACGCTGGAGCCCTGACG 701
Db 251 ArgArgLeuAlaGluPheMetGly-----CysProPheSerAlaAlaGluGluAlaAla 268
Qy 702 GAGCAGCTGCCAGCTGTGTGGACCACTGTGCAACGCTGAG-----GTG 755
Db 269 GlyAsnValAlaAlaValValGluLeuCysSerPheAspGluMetLysGlyLeuGluVal 288
Qy 744 -----GCCCTGCCC-----GTG 755
Db 289 AsnArgProGlyGlyThrAlaGlyLysTyrArgAlaMetProArgAspAlaPhePhe 308
Qy 756 GGCGCGGAGAGATTGGCTGTGGAGGACATCTTCAACCGTCTCCATGAATGAGAACTTT 815
Db 309 ArgLysGlyValAlaGlyAspTrpAlaAsnHisMetThrProGluMetAlaAlaArgLeu 328
Qy 816 GACTTGTGTATAAACAGAGATGGGAAAGTGTGACCTCACGTTT 860
Db 329 AspGlnIlePheArgAspLysLeuGlnGlyThrGlyLeuAlaPhe 343

RESULT 99
US-10-259-194A-330
; Sequence 330, Application US/10259194A
; Publication No. US20040010815A1
; GENERAL INFORMATION:
; APPLICANT: Lange, Markus B.
; APPLICANT: Ghasssemin, Majid
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Kreps, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF PLANT GENES
; FILE REFERENCE: 70029-NP
; CURRENT APPLICATION NUMBER: US/10/259,194A
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,743
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 662
; SOFTWARE: PatentList.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 330
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Oryza sativa
; US-10-259-194A-330

Alignment Scores:
Pred. No.: 3,59e-13 Length: 343
Score: 267.00 Matches: 91
Percent Similarity: 44.2% Conservative: 58
Best Local Similarity: 27.0% Mismatches: 122
```

```
Query Match: 6.1% Indels: 66
DB: 4 Gaps: 16
US-10-768-158-1 (1-2419) x US-10-259-194A-330 (1-343)
QY 9 GACGGCGGCATGCGGAGGAGCGAGCGCCGAGACCCCGAGCCCGGGGGAGTTC--- 65
D 16 AspAspGlyThrLeuAlaAlaGluAlaArgSerProLeuProLysGluGluPheGly 35
QY 66 -----GAGAGCAAGTACTTCGAG-----TTCCATGGC 92
D 36 AspleuValAlaAlaLeuProArgLysGluGlnTyrLeuAspGlyArgLeuTyrGluGly 55
QY 93 GTGGCGTCCGCCCC---TTTCGCGCGGGAAGATGGAGGAGATCGCAACTTCCCGGTG 149
D 56 PheTrpLeuProGluHisTyrAlaProGlyIleAlaPheArgArgPheThrPro 75
QY 150 CGGCGCCGAGCGTGTGATCGTCACCTACCCCAAGTCCGCGAGCTTCTCGCAGGAG 209
D 76 ArgAlaAspValLeuAlaSerTyrProLysCysGlyThrTrpLeuLysAla 95
QY 210 GTGGTCTAC---TTGGTGAGCCAGGGCGCTGACCCC---GATGAGATCGGCTTGAATGAC 263
D 96 LeuAlaPheAlaAlaMetThrArgAlaAlaTyrProAlaAspGluHisProLeuLeuArg 115
QY 264 ATCGAC-----GAGCAGTCCCGGCTCTGAGGTACCCACAGCGCGGCTTGGACATCATC 317
D 116 LeuAsnProHisAspValIleProPheValGlu-----AspValPhe 129
QY 318 AAGGAA-----CTGACCTCTCCCGCCTCATCAAGAGCCAC 353
D 130 ThrAspGlyHisGluAlaLysLeuAspMetLeuProSerProArgLeuIleAsnThrHis 149
QY 354 CTGCGCTTACGCTTTCGCGCTCTGAGGTACCCACAGCGCGGCTTGGACATCATC 401
D 150 ThrProTyrGlnLeuLeuProGluSerValValAlaGlyAspGlyGlyCysAlaLys 169
QY 402 GTCATCTATATGCTCGCAACCCCAAGGATCTGTTGTTTATTATCATGTTCCACCGC 461
D 170 ValValTyrIleCysArgAspProLysAspMetValValSerLeuTyrHisPheMetArg 189
QY 462 TCTCTGGG---ACATGAGCTACCGAGGACCTTTCAGAAATCTCGCGGAGGTTTATG 518
D 190 ArgLeuGlnProAspLeuSerLeuAlaGlyValValGluSerValAla----- 205
QY 519 AATGATAAGCTGGGCTACGGCTCTGTTTCAGCAGCTGAGGAGTTCGGGAGCACCGC 578
D 206 AspGlyThrValProPheGlyProMetTrpAspHisIleLeuGlyTyrTrpArgAlaSer 225
QY 579 ATGGAC-----TCGAACGTGCTTTTCTCAAGTATGAAGACATGCATCGGAGCTGGTG 632
D 226 ValSerArgProAspArgValLeuLeuLeuArgTyrGluAspLeuLeuArgAspGlyAla 245
QY 633 ACG-----ATGGTGGAGCAGCTGCCAGATCTCGGGGTGCTCTGTGACAGGCCAG 686
D 246 AlaGlyGluHisValArgAlaMetAlaArgPheMetGlyArgProPheSerAlaAla--- 264
QY 687 CTGGAAGCCCTGACGGAGCACTGCCACCGAGCTGTGGACAGAGTCTGCTG-----AACGCT 740
D 265 ---GluGluAlaAlaGlyAlaValAlaSerValValGluLeuCysSerPheGluArgMet 283
QY 741 GAGCGCTTCCGCTGGGCGCG-----CTGACCTCTCCCGCCTCATCAAGAGCCAC 353
D 284 LysAlaLeuGluValAsnArgArgGlyThrAlaGlySerTyrLysSerMetProArgAsp 303
QY 762 -----GGNAGAGTGGGCTGTGGAGGACATCTTCACCGTCTCCATGAAT 806
D 304 AlaPheArgLysGlyValAlaGlyAspTrpAlaAsnHisMetSerProGluThrAla 323
QY 807 GAGAAGTTTGAAGTGTATATAACAGAGATGGGAAAGTGTGACCTCACG 857
D 324 AlaArgLeuAspGlyIlePheArgGluLysPheArgGlyThrGlyLeuThr 340
```

```
RESULT 100
US-10-437-963-159374
; Sequence 159374, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 159374
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_58757C.1.pep
US-10-437-963-159374
Alignment Scores:
Pred. No.: 3 59e-13 Length: 343
Score: 267.00 Matches: 91
Percent Similarity: 44.2% Conservatives: 58
Best Local Similarity: 27.0% Mismatches: 122
Query Match: 6.1% Indels: 66
DB: 4 Gaps: 16
US-10-768-158-1 (1-2419) x US-10-437-963-159374 (1-343)
```

```
QY 9 GACGGCGGCATGCGGAGGAGCGAGCCGAGACCCCGAGCACCCCGGGGGAGTTC--- 65
D 16 AspAspGlyThrLeuAlaAlaGluAlaArgSerProLeuProLysGluGluPheGly 35
QY 66 -----GAGAGCAAGTACTTCGAG-----TTCCATGGC 92
D 36 AspleuValAlaAlaLeuProArgLysGluGlnTyrLeuAspGlyArgLeuTyrGluGly 55
QY 93 GTGGCGTCCGCCCC---TTTCGCGCGGGAAGATGGAGGAGATCGCAACTTCCCGGTG 149
D 56 PheTrpLeuProGluHisTyrAlaProGlyIleAlaPheArgArgPheThrPro 75
QY 150 CGGCGCCGAGCGTGTGATCGTCACCTACCCCAAGTCCGCGAGCTTCTGCTCGAGGAG 209
D 76 ArgAlaAspValValLeuAlaSerTyrProLysCysGlyThrTrpLeuLysAla 95
QY 210 GTGGTCTAC---TTGGTGAGCCAGGGCGCTGACCCC---GATGAGATCGGCTTGAATGAC 263
D 96 LeuAlaPheAlaAlaMetThrArgAlaAlaTyrProAlaAspGluHisProLeuLeuArg 115
QY 264 ATCGAC-----GAGCAGTCCCGGCTCTGAGGTACCCACAGCGCGGCTTGGACATCATC 317
D 116 LeuAsnProHisAspValIleProPheValGlu-----AspValPhe 129
QY 318 AAGGAA-----CTGACCTCTCCCGCCTCATCAAGAGCCAC 353
D 130 ThrAspGlyHisGluAlaLysLeuAspMetLeuProSerProArgLeuIleAsnThrHis 149
QY 354 CTGCGCTTACGCTTTCGCGCTCTGAGGTACCCACAGCGCGGCTTGGACATCATC 401
D 150 ThrProTyrGlnLeuLeuProGluSerValValAlaGlyAspGlyGlyCysLys 169
QY 402 GTCATCTATATGCTCGCAACCCCAAGGATCTGTTGTTTATTATCATGTTCCACCGC 461
D 170 ValValTyrIleCysArgAspProLysAspMetValValSerLeuTyrHisPheMetArg 189
```



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 17, 2006, 09:56:10 ; Search time 16.1 Seconds  
(without alignments)  
3945.402 Million cell updates/sec

Title: US-10-768-158-1

Perfect score: 4401

Sequence: 1 GCGACGGCGAGCGCGCGG.....aaaaaaaaaaaaaaaaaaaaa 2419

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 1301182

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-O=/absf/ABSSWEB/spool/US10768158/runat\_16052006\_163427\_5370/app\_query\_fasta\_1  
-DB-Issued\_Patents\_AA -QPMF-fastan -SUFFIX=n2p.ra -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=DIOSUM62 -TRANS=human40.cdi  
-LIST=100 -DOCALIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=100  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-HOST=absf02p -USR=US10768158 @CGN 1 1 71 @runat\_16052006\_163427\_5370 -NCPU=6  
-ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WABN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6 -FCAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Issued Patents AA:\*
- 1: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/5 COMB.pcp.\*
  - 2: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/6 COMB.pcp.\*
  - 3: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/7 COMB.pcp.\*
  - 4: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/H COMB.pcp.\*
  - 5: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/PCTUS COMB.pcp.\*
  - 6: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/RE COMB.pcp.\*
  - 7: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/backfilee1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	1530	34.8	284	2	US-09-786-240-7
2	472	10.7	295	2	US-09-854-122-21
3	456	10.4	295	2	US-09-328-174A-3
4	449	10.2	296	2	US-09-949-016-6018
5	449	10.2	296	3	US-09-792-695A-2
6	435	9.9	309	2	US-09-949-016-10328
7	434	9.9	294	1	US-08-325-562-2
8	434	9.9	294	1	US-08-437-795-2
9	430	9.8	295	2	US-08-150-133-13
10	430	9.8	295	2	US-09-150-141-13
11	430	9.8	295	2	US-09-374-493-13
12	430	9.8	295	2	US-09-786-240-7
13	430	9.8	295	2	US-09-854-122-21
14	430	9.8	295	2	US-09-328-174A-3
15	430	9.8	295	2	US-09-949-016-6018
16	430	9.8	295	2	US-09-792-695A-2
17	422	9.6	304	2	US-09-949-016-6018
18	422	9.6	304	2	US-09-792-695A-2
19	422	9.6	304	2	US-09-949-016-6018
20	422	9.6	304	2	US-09-792-695A-2
21	419.5	9.5	283	2	US-09-949-016-6018
22	419.5	9.5	283	2	US-09-792-695A-2
23	419.5	9.5	283	2	US-09-949-016-6018
24	419.5	9.5	283	2	US-09-792-695A-2
25	419.5	9.5	283	2	US-09-949-016-6018
26	419.5	9.5	283	2	US-09-792-695A-2
27	419.5	9.5	283	2	US-09-949-016-6018
28	419.5	9.5	283	2	US-09-792-695A-2
29	408.5	9.3	283	2	US-09-949-016-6018
30	408.5	9.3	283	2	US-09-792-695A-2
31	408.5	9.3	283	2	US-09-949-016-6018
32	408.5	9.3	283	2	US-09-792-695A-2
33	407	9.2	304	2	US-09-949-016-6018
34	407	9.2	304	2	US-09-792-695A-2
35	407	9.2	304	2	US-09-949-016-6018
36	407	9.2	304	2	US-09-792-695A-2
37	401	9.1	304	2	US-09-949-016-6018
38	401	9.1	304	2	US-09-792-695A-2
39	401	9.1	304	2	US-09-949-016-6018
40	401	9.1	304	2	US-09-792-695A-2
41	381	8.7	350	2	US-09-949-016-6018
42	381	8.7	350	2	US-09-792-695A-2
43	380	8.6	264	2	US-09-949-016-6018
44	333	7.6	350	1	US-08-852-481-2
45	313	7.1	60	2	US-09-513-999C-6719
46	303.5	6.9	303	2	US-09-795-926-2
47	303.5	6.9	303	2	US-10-364-774-2
48	292.5	6.6	265	2	US-09-795-926-6
49	292.5	6.6	265	2	US-10-364-774-6
50	275.5	6.3	320	2	US-09-854-122-20
51	265.5	6.0	324	2	US-09-854-122-18
52	261	5.9	302	2	US-09-854-122-19
53	254	5.8	214	2	US-09-795-926-16
54	254	5.8	214	2	US-10-364-774-16
55	212	4.8	636	2	US-09-252-991A-32429
56	208	4.7	369	2	US-09-252-991A-25394
57	199.5	4.5	263	2	US-09-252-991A-16906
58	190	4.3	282	2	US-09-252-991A-26169
59	187	4.2	186	2	US-09-795-926-12
60	187	4.2	186	2	US-10-364-774-12
61	187	4.2	331	2	US-09-854-122-17
62	187	4.2	391	2	US-09-854-122-16
63	184.5	4.2	252	2	US-09-252-991A-25346
64	183.5	4.2	481	2	US-09-949-016-9748
65	178	4.0	326	2	US-09-252-991A-30870
66	178	4.0	328	2	US-09-252-991A-21969
67	176.5	4.0	202	2	US-09-252-991A-28625
68	176.5	4.0	525	2	US-09-252-991A-23981
69	176	4.0	148	2	US-09-795-926-8
70	176	4.0	148	2	US-10-364-774-8
71	176	4.0	258	2	US-09-252-991A-22452
72	176	4.0	313	2	US-09-252-991A-27480
73	175	4.0	267	2	US-09-252-991A-28620
74	174	4.0	320	2	US-09-252-991A-24634
75	173.5	3.9	657	2	US-09-252-991A-28001
76	172.5	3.9	1706	2	US-09-252-991A-31760
77	171	3.9	335	2	US-09-252-991A-23674
78	170	3.9	418	2	US-09-252-991A-32035
79	169.5	3.9	308	2	US-09-252-991A-27249
80	167.5	3.8	697	2	US-09-252-991A-24009
81	166.5	3.8	286	2	US-09-252-991A-30343
82	166.5	3.8	335	2	US-09-252-991A-24899
83	166	3.8	244	2	US-09-252-991A-23991
84	166	3.8	437	2	US-09-252-991A-26942

c 85 165.5 3.8 306 2 US-09-252-991A-25699 Sequence 25699, A  
 c 86 165.5 3.8 511 2 US-09-252-991A-26078 Sequence 26078, A  
 c 87 164.5 3.7 218 2 US-09-252-991A-24321 Sequence 24321, A  
 c 88 164.5 3.7 315 2 US-09-252-991A-20553 Sequence 20553, A  
 c 89 164 3.7 288 2 US-09-252-991A-32807 Sequence 32807, A  
 c 90 163.5 3.7 341 2 US-09-252-991A-32424 Sequence 32424, A  
 c 91 163 3.7 233 2 US-09-252-991A-18455 Sequence 18455, A  
 c 92 163 3.7 246 2 US-09-252-991A-16586 Sequence 16586, A  
 c 93 163 3.7 248 2 US-09-252-991A-22391 Sequence 22391, A  
 c 94 162.5 3.7 663 2 US-09-252-991A-30843 Sequence 30843, A  
 c 95 161.5 3.7 247 2 US-09-252-991A-26899 Sequence 26899, A  
 c 96 161.5 3.7 272 2 US-09-252-991A-31371 Sequence 31371, A  
 c 97 160.5 3.6 333 2 US-09-252-991A-28132 Sequence 28132, A  
 c 98 160.5 3.6 663 2 US-09-252-991A-30843 Sequence 30843, A  
 c 99 160 3.6 208 2 US-09-252-991A-32166 Sequence 32166, A  
 c 100 159.5 3.6 195 2 US-09-252-991A-30009 Sequence 30009, A

## ALIGNMENTS

## RESULT 1

US-09-786-240-7  
 ; Sequence 7, Application US/09786240  
 ; Patent No. 6558935  
 ; GENERAL INFORMATION:  
 ; APPLICANT: INCYTE PHARMACEUTICALS, INC.  
 ; APPLICANT: TANG, Y. Tom  
 ; APPLICANT: CORLEY, Neil C.  
 ; APPLICANT: GUGGLER, Karl J.  
 ; APPLICANT: BAUGHN, Mariah R.  
 ; APPLICANT: LAL, Preeti  
 ; APPLICANT: YUE, Henry  
 ; APPLICANT: HILMAN, Jennifer L.  
 ; APPLICANT: AZIMZAI, Yalda  
 ; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS  
 ; FILE REFERENCE: PF-0592 PCT  
 ; CURRENT APPLICATION NUMBER: US/09/786,240  
 ; CURRENT FILING DATE: 2002-03-12  
 ; PRIOR FILING DATE: 09/150,657; unassigned; 09/186,779; unassigned; 60/133,642  
 ; PRIOR FILING DATE: 1998-09-10; 1998-11-04; 1998-11-04; 1999-05-11  
 ; NUMBER OF SEQ ID NOS: 33  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 7  
 ; LENGTH: 284  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No. 6558935 1784742CD1  
 US-09-786-240-7

Alignment Scores:  
 Pred. No.: 5.32e-142 Length: 284  
 Score: 1530.00 Matches: 284  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 34.8% Indels: 0  
 DB: 2 Gaps: 0

US-10-768-158-1 (1-2419) x US-09-786-240-7 (1-284)

Qy 21 ATGGCGGAGCGAGCGGCGAGACCCCGAGCCCGGGGGAGTTTCGAGAGCAAGTACTTC 80  
 Db 1 MetAlaGluSerGluAlaGluThrProSerThrProGluPheGluSerLysIlePhe 20  
 Qy 81 GAGTTCATGCGTGGCGGCTCCGCCCTTCTGCGCGGGAAGATGGAGGATCGCCAAAC 140  
 Db 21 GluPheHisGlyValArgLeuProPheCysArgGlyLysMetGluGluIleAlaAsn 40  
 Qy 141 TTCCGGTGGCGGCGGCGAGCGTGGATCGTACCTACCCCAAGTCGGGACAGCTTG 200  
 Db 41 PheProValArgProSerAspValTrpIleValThrProLysSerGlyThrSerLeu 60

Qy 201 CTGCAGGAGTGGTCTACTTGGTGAGCAGCGGCGCTGACCCCGATGAGATCGGCTTGATG 260  
 Db 61 LeuGlnGluValValTyrLeuValSerGlnGlyAlaAspProaspGluIleGlyLeuMet 80  
 Qy 261 AACATCGACGAGCAGCTCCCGGTCTCTGAGTACCACAGCCGGGCTTGACATCATCAAG 320  
 Db 81 AsnIleAspGluGlnLeuProValLeuGluTyrProGlnProGlyLeuAspIleIleLys 100  
 Qy 321 GAAGTACCTCTCCCGGCTCATCAAGAGCCACCTGCGCTACCGCTTCTGCGCTCTGAC 380  
 Db 101 GluLeuThrSerProArgLeuIleLysSerHisLeuProTyrArgPheLeuProSerAsp 120  
 Qy 381 CTCCACAATCGAGACTCCAAAGTCACTATATGCTCGCAACCCCAAGGATCTGGTGGTG 440  
 Db 121 LeuHisAsnGlyAspSerLysValIleTyrMetAlaArgAsnProLysAspLeuValVal 140  
 Qy 441 TCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGTACCGAGGCACCTTTCAAGAA 500  
 Db 141 SerTyrTyrGlnPheHisArgSerLeuArgThrMetSerTyrArgGlyThrPheGlnGlu 160  
 Qy 501 TTCTCCCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCTGGTTTGACAGCTGCAG 560  
 Db 161 PheCysArgArgPheMetAsnAspLysLeuGlyTyrGlySerTrpPheGluHisValGln 180  
 Qy 561 GAGTTCTGGAGCACCGCATGGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATGCAT 620  
 Db 181 GluPheTrpGluHisArgMetAspSerAsnValLeuPheLeuLysIleGluAspMetHis 200  
 Qy 621 CGGACCTCGTGACGATGGTGGAGCAGCTGCCAGATTCCTGGGGGTCTCTGTGACAAG 680  
 Db 201 ArgAspLeuValThrMetValGluGlnLeuAlaArgPheLeuGlyValSerCysAspLys 220  
 Qy 681 GCCAGCTGGAAGCCCTGACGGAGCACTGCCACAGCTGTGGACAGTGTGCAACGCT 740  
 Db 221 AlaGlnLeuGluAlaLeuThrGluHisCysHisGlnLeuValAspGlnCysCysAsnAla 240  
 Qy 741 GAGCCCTGCGCGTGGCGGGAAGAGTTGGGTGTGGAGGACATCTTCACCGTCTCC 800  
 Db 241 GluAlaLeuProValGlyArgGlyArgValGlyLeuTrpLysAspIlePheThrValSer 260  
 Qy 801 ATGAATCAGAAGTTGACTTGGTGTATATAACAGAAGATGGAAAGTGTGACCTCAGCTTT 860  
 Db 261 MetAsnGluLysPheAspLeuValTyrLysGlnLysMetGlyLysCysAspLeuThrPhe 280  
 Qy 861 GACTTTTATTATA 872  
 Db 281 AspPheTyrLeu 284

RESULT 2  
 US-09-854-122-21  
 ; Sequence 21, Application US/09854122  
 ; Patent No. 6841718  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ALBERTE, RANDALL S.  
 ; APPLICANT: SMITH, ROBERT  
 ; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA  
 ; FILE REFERENCE: PHA-007.01  
 ; CURRENT APPLICATION NUMBER: US/09/854,122  
 ; CURRENT FILING DATE: 2001-09-10  
 ; PRIOR FILING DATE: 60/202,529  
 ; PRIOR FILING DATE: 2000-05-10  
 ; NUMBER OF SEQ ID NOS: 51  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 21  
 ; LENGTH: 295  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-854-122-21

Alignment Scores:  
 Pred. No.: 1.62e-37 Length: 295  
 Score: 472.00 Matches: 100  
 Percent Similarity: 54.5% Conservative: 52

```

; CURRENT FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 09/328,174
; PRIOR FILING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 295
; TYPE: PRT
; ORGANISM: H. sapiens
US-09-328-174A-3

Alignment Scores:
Pred. No.: 6,17e-36 Length: 295
Score: 456.00 Matches: 94
Percent Similarity: 54.5% Conservative: 58
Best Local Similarity: 33.7% Mismatches: 101
Query Match: 10.4% Indels: 26
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x US-09-328-174A-3 (1-295)
Qy 90 GCGGTGGGTGCGCGCCCTTCTGCCGCGGAAGATGGAGGAGATGCCAACTTCCCGTG 149
    ||||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 17 GlyValProLeuIleLysTyrPheAlaGluAlaLeuGlyProLeuGlnSerPheGlnAla 36
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 150 CGCCCCAGCAGCTGTGGATCGTCACCTACCCCAAGTCCGGCAGCAGCTGTGCGAGGAG 209
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 37 ArgProAspLeuLeuIleSerThrTyrProLysSerGlyThrTrpValSerGln 56
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 210 GTGCTCTACTTGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGAC 269
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 57 IleLeuAspMetIleTyrGlnGlyGlyAspLeuGluLysCysHisArgAlaProIlePhe 76
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 270 GACGAGCTCCGCTCTGGAGTACCCACAGCCG-----GGCTGGACATCATC 317
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 77 ValArgValProPheLeuGluValAsnAspProGlyGluProSerGlyLeuGluThrLeu 96
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 318 AAGGAAGTACCTCTCCCGCCTCATCAAGAGCCACCTGCGCTTCTCCCTTCTGCCCTCT 377
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 97 LysAsnThrProAlaProArgLeuLeuLysThrHisLeuProLeuAlaLeuLeuProGln 116
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 378 GACCTCCACATGGAGACTCCAAGGTCACTATATGGCTCGCAACCCCAAGGATCTGGTG 437
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 117 ThrLeuLeuAspGlnLysValValTyrValAlaArgAsnProLysAspValAla 136
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 438 GTGCTTATATCAGTTCCACCGCTCTCTCGGACCATGAGTACCGAGGCACCTTTCAA 497
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 137 ValSerTyrTyrHisPheHisArgMetGluLysAlaHisProGluProGlyThrTrpAsp 156
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 498 GAATTCGCGGAGTTTATGAATGATAGCTAGGCTACGGCTCGCTCTGTTGAGCAGCTG 557
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 157 SerPheLeuGluLysPheMetAlaGlyValSerTyrGlySerTrpTyrGlnHisVal 176
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 558 CAGGAGTTCTGGAGCACCAGCATGAGTGGGCTGTGGAAGGACATCTTCCCGTCTCCATG 617
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 177 GlnGluTrpTrpGluLeuSerArgThrHisProValLeuTyrLeuPheTyrGlnAspMet 196
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 618 CATCGGACCTGTGACGATGTGGAGCTGCGCAGATTCCTGGGGGTGTCGTGAC 677
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 197 LysGluAsnProLysArgGluIleGlnLysIleLeuGluPheValGlyArgSerLeuPro 216
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 678 AAGCCCGAGCTGGAAGCCCTGACCGAGCAGCTGCTGCAAGCTGAGGCC 746
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 237 MetThrAsnTyrThrThrValProGlnGluLeuMetAspHisSerIleSer----- 253
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 747 CTGCCC---GTGGCCCGGGAAGAGTTGGGCTGTGGAAGGACATCTTCCCGTCTCCATG 803
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 254 --ProPheMetArgLysGlyMetAlaGlyAspTrpLysThrThrPheThrValAlaGln 272
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 804 AATGAGAAGTTTGCATCTGTTGATATAACAGAGATGGGAAAGTGATGACCTACAGTTT 860
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291
    ||||| |||| : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
US-09-328-174A-3
; Sequence 3, Application US/09328174A
; Patent No. 6448003
; GENERAL INFORMATION:
; APPLICANT: Guida, Marco
; APPLICANT: Kurth, Janice
; TITLE OF INVENTION: Genotyping Human Phenol Sulfotransferase
; FILE REFERENCE: 4389-6 (formerly SEQ-16P)
; CURRENT APPLICATION NUMBER: US/09/328,174A

```

Qy 804 AATGAGAAGTTTGACTTGGTGATATAACAGAGAGATGGGAAAGTGTGACCTCACGTTT 860  
 |||||:::|||||  
 Db 273 AsnGluArgPheAspAlaAspTyrAlaGluLysMetAlaGlyCysSerLeuSerPhe 291  
 |||||:::|||||  
 RESULT 4  
 US-09-949-016-6018  
 ; Sequence 6018, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: C0001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; PRIOR FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 6018  
 ; LENGTH: 296  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 US-09-949-016-6018

Alignment Scores:  
 Pred. No.: 3,04e-35 Length: 296  
 Score: 449.00 Matches: 94  
 Percent Similarity: 54.6% Conservative: 61  
 Best Local Similarity: 33.1% Mismatches: 105  
 Query Match: 10.2% Indels: 24  
 DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x US-09-949-016-6018 (1-296)

Qy 72 AAGTACTTCGAGTTCATGGCTGGCTGCCGCTCTCTCCCGGGAGATGGAGGAG 131  
 |||||:::|||||  
 Db 12 LysLeuLysGluValGluGlyThrLeuLeuGlnProAlaThrValAspAsnTrpSerGln 31  
 |||||:::|||||

Qy 132 ATCCGCCAATTCCTCCGCTGCCGCGCCAGACGCTGGATCGTCACCTACCCCAAGTCCGGC 191  
 |||||:::|||||  
 Db 32 IleGlnSerPheGluAlaLysProAspAspLeuLeuLeuCysThrTyrProLysAlaGly 51  
 |||||:::|||||

Qy 192 ACCAGCTTGCTGCAGGAGTGTCTACTTGGTGAGCCAGCGCTGACCCCGATGAGATC 251  
 |||||:::|||||  
 Db 52 ThrThrTrpIleGlnGluValAspMetIleGlnAsnGlyAspValGluLysCys 71  
 |||||:::|||||

Qy 252 GGCTTGATGAACATCGACGAGCTCCCGCTCTCGAGTAC-----CCACAGCGC 302  
 |||||:::|||||  
 Db 72 GlnArgAlaIleGlnHisProPheIleGluTrpAlaArgProProGlnPro 91  
 |||||:::|||||

Qy 303 ---GGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACTGCC 359  
 |||||:::|||||  
 Db 92 SerGlyValGluLysAlaLysAlaMetProSerProArgIleLeuLysThrHisLeuSer 111  
 |||||:::|||||

Qy 360 TACCGCTTCTGCCCTCTGACCTCCCAATGGAGACTCCCAAGGTCATCTATATGGCTGC 419  
 |||||:::|||||  
 Db 112 ThrGlnLeuLeuProProSerPheTrpGluAsnAsnCysLysPheLeuTyrValAlaArg 131  
 |||||:::|||||

Qy 420 AACCCCAAGGATCTGGTGTCTTATATCAGTTTCACCGCTCTCTCGGACCATGAGC 479  
 |||||:::|||||  
 Db 132 AsnAlaLysAspCysMetValSerTyrTyrHisPheGlnArgMetAsnHisMetLeuPro 151  
 |||||:::|||||

Qy 480 TACCGAGCCACTTCAAGAACTTCGCGGAGGTTTATGAATGATAAGCTGGGTACGGC 539  
 |||||:::|||||  
 Db 152 AspProGlyThrTrpGluLysPheGluThrPheIleAsnGlyLysValValTrpGly 171  
 |||||:::|||||

Qy 540 TCCTGGTTGAGCAGCTGCAGGAGTCTTGGGAGCACCGCATCGGACTCGCACTGCTTTT 599  
 |||||:::|||||

Db 172 SerTrpPheAspHisValLysGlyTrpTrpGluMetLysAspArgHisGlnIleLeuPhe 191  
 |||||:::|||||  
 Qy 600 CTCGAAGTATCAAGACATGCGGACCTGGTACGATGCTGGAGCAGCTGCCAGATTC 659  
 |||||:::|||||  
 Db 192 LeuPheTyrGluAspIleLysArgAspProLysHisGluIleArgLysValMetGlnPhe 211  
 |||||:::|||||  
 Qy 660 CTGGGGGTGCTCTGTGACAAAGGCCAGCTGGAGCCCTGACGGAGCAC-----707  
 |||||:::|||||  
 Db 212 MetGlyLysLysValAspGluThrValLeuAspLysIleValGlnGluThrSerPheGlu 231  
 |||||:::|||||  
 Qy 708 -----TGCCACCAGCTGGTGAGCAC 728  
 |||||:::|||||  
 Db 232 LysMetLysGluAsnProMetThrAsnArgSerThrValSerLysSerIleLeuAspGln 251  
 |||||:::|||||  
 Qy 729 TGCTGCAACGCTGAGGCCCTGCCGCTGGCGGGGAGAGCTTGGCTGTGGAAGACATC 788  
 |||||:::|||||  
 Db 252 SerIleSerSerPhe-----MetArgLysGlyThrValGlyAspTrpLysAsnHis 268  
 |||||:::|||||  
 Qy 789 TTCACCGCTCTCCATGAATGAGAAAGTTTGACTTGGTGATATAACAGAGATGGGAAAGTGT 848  
 |||||:::|||||  
 Db 269 PheThrValAlaGlnAsnGluArgPheAspGluIleTyrArgArgLysMetGluGlyThr 288  
 |||||:::|||||  
 Qy 849 GACCTCACGTTT 860  
 |||||:::|||||  
 Db 289 SerIleAsnPhe 292  
 |||||:::|||||

## RESULT 5

US-09-792-695A-2  
 ; Sequence 2, Application US/09792695A  
 ; Patent No. 7026163  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Freimuth, Robert R.  
 ; APPLICANT: Weinshilboum, Richard M  
 ; APPLICANT: Wieben, Eric D.  
 ; TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS  
 ; FILE REFERENCE: 07039-273001  
 ; CURRENT APPLICATION NUMBER: US/09/792,695A  
 ; CURRENT FILING DATE: 2001-02-23  
 ; NUMBER OF SEQ ID NOS: 37  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 296  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-792-695A-2

Alignment Scores:  
 Pred. No.: 3,04e-35 Length: 296  
 Score: 449.00 Matches: 94  
 Percent Similarity: 54.6% Conservative: 61  
 Best Local Similarity: 33.1% Mismatches: 105  
 Query Match: 10.2% Indels: 24  
 DB: 3 Gaps: 4

US-10-768-158-1 (1-2419) x US-09-792-695A-2 (1-296)

Qy 72 AAGTACTTCGAGTTCATGGCTGGCTGCCGCTCTCTCCCGGGAGATGGAGGAG 131  
 |||||:::|||||  
 Db 12 LysLeuLysGluValGluGlyThrLeuLeuGlnProAlaThrValAspAsnTrpSerGln 31  
 |||||:::|||||

Qy 132 ATCCGCCAATTCCTCCGCTGCCGCGCCAGACGCTGGATCGTCACCTACCCCAAGTCCGGC 191  
 |||||:::|||||  
 Db 32 IleGlnSerPheGluAlaLysProAspAspLeuLeuLeuCysThrTyrProLysAlaGly 51  
 |||||:::|||||

Qy 192 ACCAGCTTGCTGCAGGAGTGTCTACTTGGTGAGCCAGCGCTGACCCCGATGAGATC 251  
 |||||:::|||||  
 Db 52 ThrThrTrpIleGlnGluValAspMetIleGlnAsnGlyAspValGluLysCys 71  
 |||||:::|||||

Qy 252 GGCTTGATGAACATCGACGAGCTCCCGCTCTCGAGTAC-----CCACAGCGC 302  
 |||||:::|||||  
 Db 72 GlnArgAlaIleGlnHisProPheIleGluTrpAlaArgProProGlnPro 91  
 |||||:::|||||

Qy 303 ---GGCTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACTGCC 359  
 |||||:::|||||



```

Db      92  SerGlyValGluLysAlaLysAlaMetProSerProArgIleLeuLysThrHisLeuSer 111
Qy      360  TACCGCTTTCTGCGCTCTGAGCTCCACCAATGGAGAGTTCATATATATGCTCGC 419
Db      112  ThrGlnLeuLeuProSerPheTrpGluAsnAsnCysLysPheLeuTyValAlaArg 131
Qy      420  AACCCCAAGGATCTGGTGTCTTATATATACAGTTTCCACCGCTCTCTCGGACCATGAGC 479
Db      132  AsnAlaLysAspCysMetValSerTyTrHisPheGlnArgMetAsnHisMetLeuPro 151
Qy      480  TACCGAGCCACTTTCAAGATTCTCGCGAGGTTTATGAATGATAAGCTGGCTACGGC 539
Db      152  AspProGlyThrTrpGluGlyTrPheGluThrPheIleAsnGlyLysValValTrpGly 171
Qy      540  TCCTGGTTGAGCAGCTGCAGAGTTCTGGGAGCACCGCATGGACTCGAAGCTGCTTTT 599
Db      172  SerTrpPheAspHisValLysGlyTrpTrpGluMetLysAspArgHisGlnIleLeuPhe 191
Qy      600  CTCAGTATGAACATCATCGGACCTGGTGGACGATGGTGGAGCAGCTGGCCAGATTC 659
Db      192  LeuPheTyGluAspIleLysArgAspProLysHisGluIleArgLysValMetGlnPhe 211
Qy      660  CTGGGGGTCTCTGTGACAGCCCGAGCTGGAAGCCCTGACGGAGCAC----- 707
Db      212  MetGlyLysLysValAspGluThrValLeuAspLysIleValGlnGluThrSerPheGlu 231
Qy      708  -----TGCCACCACTGGTGACCCAG 728
Db      232  LysMetLysGluAsnProMetThrAsnArgSerThrValSerLysSerIleLeuAspGln 251
Qy      729  TGCTGCAACGTGAGCCCTCCCGTGGCGGGAGAGATTGGCTGTGGAAGGACATC 788
Db      252  SerIleSerPhe-----MetArgLysGlyThrValGlyAspTrpLysAsnHis 268
Qy      789  TTCACCTCTCCATGAATGAGAAGTTTCACCTGGTGTGTATTAACAGAGATGGGAAGTGT 848
Db      269  PheThrValAlaGlnAsnGluArgPheAspGluIleTyArgArgLysMetGluGlyThr 288
Qy      849  GACCTCAGCTTT 860
Db      289  SerIleAsnPhe 292

RESULT 6
US-09-949-016-10328
; Sequence 10328, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10328
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10328

Alignment Scores:
Pred No.:          7,52e-34          Length:          309
Score:             435.00           Matches:          96
Percent Similarity: 52.7%           Conservative:     59
Best Local Similarity: 32.7%           Mismatches:       121

```

```

Query Match:      9.9%      Indels:      18
DB:                2        Gaps:        3

US-10-768-158-1 (1-2419) x US-09-949-016-10328 (1-309)

Qy      33  GAGCCGAGACCCCGACAGCACCCTGGGGAGTTCAGAGCAAGTACTTCGAGTCCATGGC 92
Db      12  GlnCysThrThrMetAsnSerGluLeuAspTyTrpGluLysPheGluGluValHisGly 31
Qy      93  GTGGCGTGGCGCCCTCTTCTGCGCGGGAAGATGAGAGATCGCAACTTCCCGGTGGG 152
Db      32  IleLeuMetTyTrpLysAspPheValLysTyTrpAspAsnValGluAlaPheGlnAlaArg 51
Qy      153  CCCAGCAGCGTGTGGATCGTCACTACCCCAAGTCCCGCACAGCTTCGACGAGGTG 212
Db      52  ProAspAspLeuValIleAlaThrTyProLysSerGlyThrTrpValSerGluIle 71
Qy      213  GTCTACTTGGTGAGCCAGGCGCTGACCCCGATGAGATCGGCTTCATGAATCAACGACG 272
Db      72  ValTyMetIleTyLysGluGlyAspValGluLysCysLysGluAspValIlePheAsn 91
Qy      273  CAGCTCCCGGTCTCTGGAGTACCCACAGCG-----GGCCTGGACATCATCAAG 320
Db      92  ArgIleProPheLeuGluCysArgLysGluAsnLeuMetAsnGlyValLysGlnLeuAsp 111
Qy      321  GAACGTACTCTCCCGCTCATCAAGACCCAGCTGCTGCTTCTGCTTCTGCTTCTGCTG 380
Db      112  GluMetAsnSerProArgIleValLysThrHisLeuProGluLeuLeuProAlaSer 131
Qy      381  CTCACAAATGAGACTCCAGGTCTATATATGCTGCAACCCCAAGGATCTGGTGTG 440
Db      132  PheTrpGluLysAspCysLysIleIleTyLysCysArgAsnAlaLysAspValAlaVal 151
Qy      441  TCTTATTATCATGTTCCACCGCTCTCTCGGACCATGAGTACCGAGGACCTTCAAGAA 500
Db      152  SerPheTyTrpPhePheLeuMetValAlaGlyHisProAsnProGlySerPheProGlu 171
Qy      501  TTCTGCCGAGGTTTATGAATGATAAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 560
Db      172  PheValGluLysPheMetGlnGlyGlnValProTyTrpGlySerTrpTyTrpLysHisVal 191
Qy      561  GAGTCTTGGGAGCACCGCATGGACTCGAACGTGCTTTTCTCAAGTATGAAGACATGAT 620
Db      192  SerTrpTrpGluLysGlyLysSerProArgValLeuPheLeuPheTyTrpGluAspLeu 211
Qy      621  CGGACCTGTGTGACGATGGTGGAGCAGCTGCCAGATTCTGGGGGTCTCTGTGTGACA 680
Db      212  GluAspIleArgLysGluValIleLysLeuIleHisPheLeuGluArgLysPheSerGlu 231
Qy      681  GCCCAGCTGGAAGCCCTGAGCGAGCATCTGCCAC----- 713
Db      232  GluLeuValAspArgIleIleHisThrSerPheGlnGluMetLysAsnAsnProSer 251
Qy      714  -----CAGCTGGTGGACCATGCTGCAACCGCTGAGGCCCTGCCCC-----GTGGGC 758
Db      252  ThrAsnTyTrpThrThrLeuProAspGluIleMetAsnGlnLysLeuSerProPheMetArg 271
Qy      759  CGGGGAAGAGTTGGCTGTGGAAGGACATCTTCAACCGTCTCCATGAATGAGATTGAC 818
Db      272  LysGlyIleThrGlyAspTrpLysAsnHisPheThrValAlaLeuAsnGluLysPheAsp 291
Qy      819  TTGGTGTATAACAGAAAGATGGGAAGTGTGACCTCACGTTT 860
Db      292  LysHisTyTrpGluGlnGlnMetLysGluSerThrLeuLysPhe 305

RESULT 7
US-08-325-562-2
; Sequence 2, Application US/08325562
; Patent No. 5714594
; GENERAL INFORMATION:
; APPLICANT: Weinshtilboun, Richard M.
; APPLICANT: Aksoy, Ibrahim A.
; APPLICANT: Wood, Thomas C.

```





```
Db 227 ThrSerPheGlnGluMetLysAsnProSerThrAsnTyrThrMetMetProGluGlu 246
Qy 729 TGCTGCAACGCTGAGCCCTGCC---GTGGCCGGGGAAGAGTTGGCTGTGAAGAC 785
Db 247 MetMetAsnGlnLysValSerProPheMetArgLysGlyIleGlyAspTrpLysAsn 266
Qy 786 ATCTTACCGCTCCATGAATGAGAAGTTTACATTGGTGTATAACAGAACATGGAAAG 845
Db 267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnMetLysAsp 286
Qy 846 TGTGACCTCACGTTT 860
Db 287 CysThrValLysPhe 291

RESULT 10
US-09-150-141-13
; Sequence 13, Application US/09150141B
; Patent No. 6071732
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES, NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 5820.546
; CURRENT APPLICATION NUMBER: US/09/150.141B
; CURRENT FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 5.1 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-150-141-13

Alignment Scores:
Pred. No.: 2,28e-33 Length: 295
Score: 430.00 Matches: 94
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 33.0% Mismatches: 116
Query Match: 9.8% Indels: 18
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x US-09-150-141-13 (1-295)
Qy 60 GAGTTCGAGACAAAGTACTTCCAGTTTCATCGCGCTGCCCGCTTCGCCGGGG 119
Db 7 GluTyrTyrGluValPheGlyGluPheArgGlyValLeuMetAspLysArgPheThrLys 26
Qy 120 AAGATGGAGGAGATCGCCAACTTCCCGTGGCGCCCGACGAGCTGTGGATCGTACCTAC 179
Db 27 TyrTrpGluAspValGluMetPheLeuAlaArgProAspAspLeuValIleAlaThrTyr 46
Qy 180 CCCAAGTCCCGCACCAGCTGTGTCGAGGAGTGTCTACTTGGTGAGCGCGCGCTGCAC 239
Db 47 ProLysSerGlyThrThrTrpLysSerGluValValTyrMetIleTyrLysGluGlyAsp 66
Qy 240 CCCGATGAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGTCTCGAGTACCCACAG 299
Db 67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86
Qy 300 CCG-----GGCTGGACATCATCAGNACTGACTCTCCCGCTCATCAG 347
Db 87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleValLys 106
Qy 348 AGCCACCTGCCTACCGCTTCTGCGCTCTGACCTCCACAAATGGAGATCCCAAGGTATC 407
Db 107 ThrHisLeuProProLysValLeuProAlaSerPheThrGluLysAsnCysLysMetIle 126
Qy 408 TATATGGCTCGCAACCCCAAGATCTGGTGTCTTATATATCATGTTCCACCGCTCTCTG 467
Db 127 TyrLeuCysArgAsnAlaLysAspValAlaValSerTyrTyrTyrPheLeuLeuMetIle 146
```

```
Qy 468 CGGACCATGAGCTACCGAGCACCTTTCAAGAATTTCTCCGAGGTTTATGATGATAAG 527
Db 147 ThrSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGlnGlyGln 166
Qy 528 CTGGGCTACGCTCCTCGTTTGGACACGTGCAGAGTTCTGGGAGCACCCGATCGACTCG 587
Db 167 ValProTyrGlySerTrpTyrAspHisValLysAlaTrpTrpGluLysSerLysAsnSer 186
Qy 588 AACCTGCTTTTTCACAGTATGAAGACATGATCGGACCTGGTGAGCATGTTGGAGGAC 647
Db 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgGluValValLys 206
Qy 648 CTGGCCAGATTCTCGGGGTCTCTGTGACAAGCCCGAGCTGGAGCCCTCACCGAGCAC 707
Db 207 LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleGlnHis 226
Qy 708 TGCCAC-----CAGCTGGTGGACGAC 728
Db 227 ThrSerPheGlnGluMetLysAsnProSerThrAsnTyrThrMetMetProGluGlu 246
Qy 729 TGCTGCAACGCTGAGCCCTGCC---GTGGCCGGGGAAGAGTTGGCTGTGAAGAC 785
Db 247 MetMetAsnGlnLysValSerProPheMetArgLysGlyIleGlyAspTrpLysAsn 266
Qy 786 ATCTTACCGCTCTCCATGAATGAGAAGTTTACATTGGTGTATAACAGAACATGGAAAG 845
Db 267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnMetLysAsp 286
Qy 846 TGTGACCTCACGTTT 860
Db 287 CysThrValLysPhe 291

RESULT 11
US-09-374-493-13
; Sequence 13, Application US/09374493
; Patent No. 6204016
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5820.546
; CURRENT APPLICATION NUMBER: US/09/374.493
; CURRENT FILING DATE: 1999-08-13
; EARLIER APPLICATION NUMBER: 09/150.133
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 60/072.994
; EARLIER FILING DATE: 1998-01-29
; EARLIER APPLICATION NUMBER: PCT/US99/16750
; EARLIER FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-374-493-13

Alignment Scores:
Pred. No.: 2,28e-33 Length: 295
Score: 430.00 Matches: 94
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 33.0% Mismatches: 116
Query Match: 9.8% Indels: 18
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x US-09-374-493-13 (1-295)
Qy 60 GAGTTCGAGACAAAGTACTTCCAGTTTCATCGCGCTGCCCGCTTCGCCGGGG 119
Db 7 GluTyrTyrGluValPheGlyGluPheArgGlyValLeuMetAspLysArgPheThrLys 26
```

```
Qy 120 AAGATGGAGGAGATCGCAACTTCCCGGTGGCGCCAGCGACGTGTGGATCGTCACCTAC 179
Db 120 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 27 TyrTrpGluAspValGluMetPheLeuAlaValPheAspGluValLeuAlaThrTyr 46
Db 27 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 180 CCCAAGTCGGCAGCAGCTTCTGCGAGGAGTGTCTACTTGGTGGCAGCCGCGCTGAC 239
Db 180 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 47 ProlYserGlyThrTrpPheLeuValTyrMetIleValTyrPheLeuGluGlyAsp 66
Db 47 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 240 CCCGATGAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGGTCTCGAGTACCCACAG 299
Db 240 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86
Db 67 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 300 CCG-----GGCTGGACATCATCAAGAACTGACCTTCCCGCTCATCAAG 347
Db 300 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 87 GluAspLeuIleAsnGlyLysGlnLeuLysGluLysGluSerProArgIleValLys 106
Db 87 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 348 AGCCACCTGCGCTACCGGCTTCTGCGCTCTGACCTCCACAAATGGAGACTCCAAGTCTATC 407
Db 348 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 107 ThrHisLeuProLysValLeuProLysPheTrpGluLysAsnCysLysMetIle 126
Db 107 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 408 TATATGGCTCGCAACCCCAAGGATCTGGTGTGTCTTATTATCATGTTCCACCGCTCTCTG 467
Db 408 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 127 TyrLeuCysArgAsnAlaLysAspValAlaValSerTyrTyrPheLeuLeuMetIle 146
Db 127 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 468 CGGACCATGAGCTACCGAGGCACCTTCAAGAATTCGCGGAGGTTTATGAATGATAAG 527
Db 468 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 147 ThrSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGlnGlyGln 166
Db 147 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 528 CTGGCTACGCTCCCTGGTTTGACGACGTCTGGAGAGTTCTGGGAGCACCGCATCGACTCG 587
Db 528 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 167 ValProTyrGlySerTrpTyrAspHisValTyrPheValTyrPheLeuLysSerLysAsnSer 186
Db 167 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 588 AACGTGTCTTCTCAAGATGAAGACATGCATCGGACCTGGTGTGACGATGGTGGAGCAG 647
Db 588 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgGluValValLys 206
Db 187 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 648 CTGCCACAGATTCCTGGGGGTCTCTGTGCAAGGCCCGCAGCTGGAAGCCCTGACGAGCAG 707
Db 648 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 207 LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleGlnHis 226
Db 207 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 708 TGCCAC-----CAGCTGGTGGACCCAG 728
Db 708 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 227 ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrMetProGluGlu 246
Db 227 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 729 TGCTGCAACGCTGAGGCCCTGCCCC---GTGGCGCGGGAAGAGTGGCGCTGTGGAAGGAC 785
Db 729 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 247 MetMetAsnGlnLysValSerProPheMetArgLysGlyIleIleGlyAspTrpLysAsn 266
Db 247 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 786 ATCTTACCGTCTCCATGAATGAGAAGTTTGACTTGGTGTATATAACAGAGATGGGAAG 845
Db 786 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnMetLysAsp 286
Db 267 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 846 TGTGACCTCAGCTTT 860
Db 846 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 287 CysThrValLysPhe 291
Db 287 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
US-09-374-824-13
; Sequence 13, Application US/09374824
; Patent No. 6207414
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5820.547
; CURRENT APPLICATION NUMBER: US/09/374,824
; CURRENT FILING DATE: 1999-08-13
; EARLIER APPLICATION NUMBER: 09/150,133
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 60/072,994
; EARLIER FILING DATE: 1998-01-29
; EARLIER APPLICATION NUMBER: PCT/US99/16750
; EARLIER FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
```

```
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-374-824-13

Alignment Scores:
Pred No.: 2,28e-33 Length: 295
Score: 430.00 Matches: 94
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 33.0% Mismatches: 116
Query Match: 9.8% Indels: 18
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x US-09-374-824-13 (1-295)
Qy 60 GAGTTCGAGACAAAGTACTTCGAGTTCATGGCTGGCGTCCGCCCTTCTCGCGCGG 119
Db 60 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 7 GluTyrTyrGluValPheGlyGluPheArgGlyValLeuMetAspLysArgPheThrLys 26
Db 7 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 120 AAGATGAGGAGATCGCAACTTCCCGGTGGCGCCAGCGACGTGTGGATCGTCACCTAC 179
Db 120 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 27 TyrTrpGluAspValGluMetPheLeuAlaArgProAspAspLeuValIleAlaThrTyr 46
Db 27 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 180 CCCAAGTCGGCAGCAGCTTGTGTCAGAGGTGTCTACTTGGTGGAGCCGCGCTGAC 239
Db 180 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 47 ProlYserGlyThrTrpPheLeuValTyrMetIleValTyrPheLeuGluGlyAsp 66
Db 47 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 240 CCCGATGAGATCGGCTTGATGAACATCGACGAGCAGCTCCCGGTCTCGAGTACCCACAG 299
Db 240 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86
Db 67 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 300 CCG-----GGCTGGACATCATCAAGAACTGACCTTCCCGCTCATCAAG 347
Db 300 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleValLys 106
Db 87 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 348 AGCCACCTGCGCTACCGGCTTCTGCGCTCTGACCTCCACAAATGGAGACTCCAAGTCTATC 407
Db 348 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 107 ThrHisLeuProLysValLeuProLysPheTrpGluLysAsnCysLysMetIle 126
Db 107 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 408 TATATGGCTCGCAACCCCAAGGATCTGGTGTGTCTTATTATCATGTTCCACCGCTCTCTG 467
Db 408 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 127 TyrLeuCysArgAsnAlaLysAspValAlaValSerTyrTyrPheLeuLeuMetIle 146
Db 127 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 468 CGGACCATGAGCTACCGAGGCACCTTCAAGAATTCGCGGAGGTTTATGAATGATAAG 527
Db 468 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 147 ThrSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGlnGlyGln 166
Db 147 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 528 CTGGCTACGCTCCCTGGTTTGACGACGTCTGGAGAGTTCTGGGAGCACCGCATCGACTCG 587
Db 528 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 167 ValProTyrGlySerTrpTyrAspHisValTyrPheValTyrPheLeuLysSerLysAsnSer 186
Db 167 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 588 AACGTGTCTTCTCAAGATGAAGACATGCATCGGACCTGGTGTGACGATGGTGGAGCAG 647
Db 588 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgGluValValLys 206
Db 187 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 648 CTGCCACAGATTCCTGGGGGTCTCTGTGCAAGGCCCGCAGCTGGAAGCCCTGACGAGCAG 707
Db 648 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 207 LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleGlnHis 226
Db 207 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 708 TGCCAC-----CAGCTGGTGGACCCAG 728
Db 708 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 227 ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrMetProGluGlu 246
Db 227 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 729 TGCTGCAACGCTGAGGCCCTGCCCC---GTGGCGCGGGAAGAGTGGCGCTGTGGAAGGAC 785
Db 729 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 247 MetMetAsnGlnLysValSerProPheMetArgLysGlyIleIleGlyAspTrpLysAsn 266
Db 247 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 786 ATCTTACCGTCTCCATGAATGAGAAGTTTGACTTGGTGTATATAACAGAGATGGGAAG 845
Db 786 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnMetLysAsp 286
Db 267 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 846 TGTGACCTCAGCTTT 860
Db 846 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 287 CysThrValLysPhe 291
Db 287 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnGlnMetLysAsp 286  
Qy 846 TGTGACCTCAGTTT 860  
||| :|||  
Db 287 CysThrValLysPhe 291  
||| :|||  
RESULT 13  
US-09-374-492-13  
; Sequence 13, Application US/09374492  
; Patent No. 6207432  
; GENERAL INFORMATION:  
; APPLICANT: The Board of Regents of the University of Oklahoma  
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF  
; FILE REFERENCE: 5820.545  
; CURRENT APPLICATION NUMBER: US/09/374,492  
; CURRENT FILING DATE: 1999-08-13  
; EARLIER APPLICATION NUMBER: 09/150,141  
; EARLIER FILING DATE: 1998-09-09  
; EARLIER APPLICATION NUMBER: 60/072,994  
; EARLIER FILING DATE: 1998-01-29  
; EARLIER APPLICATION NUMBER: PCT/US99/16750  
; EARLIER FILING DATE: 1999-07-23  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)  
; SEQ ID NO 13  
; LENGTH: 295  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; FEATURE:  
; OTHER INFORMATION: estrogen sulfotransferase  
US-09-374-492-13  
Alignment Scores:  
Pred. No.: 2,28e-33 Length: 295  
Score: 430.00 Matches: 94  
Percent Similarity: 53.0% Conservative: 57  
Best Local Similarity: 33.0% Mismatches: 116  
Query Match: 9.8% Indels: 18  
Gaps: 3  
US-10-768-158-1 (1-2419) x US-09-374-492-13 (1-295)  
Qy 60 GAGTTCGAGCAAGTACTTCGAGTTCCATCGCTGGCGTCCGCCCTTCGCGCGG 119  
||| :|||  
Db 7 GluTyrGluValPheGlyGluPheArgGlyValLeuMetAspLysArgPheThrLys 26  
||| :|||  
Qy 120 AAGATGGAGGAGATCGCAACTTCCCGTGGCGGCCAGCGAGTGTGATCGTCACCTAC 179  
||| :|||  
Db 27 TyrTrpGluAspValGluMetPheLeuAlaArgProAspAspLeuValIleAlaThrTyr 46  
||| :|||  
Qy 180 CCCAAGTCCGACACGAGTTCCTCGCAGGAGTGTCTACTTGTGAGCCGCGCTGAC 239  
||| :|||  
Db 47 ProlYsSerGlyThrThrTrpIleSerGluValValTyrMetIleTyrLysGluGlyAsp 66  
||| :|||  
Qy 240 CCCGATGATCGGCTTCATCAACATCGACGAGCTCCCGTCCGCTCGAGTACCCACAG 299  
||| :|||  
Db 67 ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn 86  
||| :|||  
Qy 300 CCG-----GGCTCGACATCATCAAGAACTGACCTCTCCCGCGCTCATCAAG 347  
||| :|||  
Db 87 GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleValLys 106  
||| :|||  
Qy 348 AGCCACCTGCGCTACCGCTTCTCGCTCTGACCTCCACATGAGAGACTCCAGGTCTATC 407  
||| :|||  
Db 107 ThrHisLeuProLysValLeuProAlaSerPheTrpGluLysAsnCysLysMetIle 126  
||| :|||  
Qy 408 TATATGGCTGCAACCCCAAGGATCTGGTGTCTTATTATCATGTTCCACCGCTCTCTG 467  
||| :|||  
Db 127 TyrLeuCysArgAsnAlaLysAspValAlaValSerTyrTyrPheLeuLeuMetIle 146  
||| :|||  
Qy 468 CGGACCATGAGTACCGGACGACCTTTCAGAATTCGCCGAGGTTTATGAATGATAG 527  
||| :|||  
Db 147 ThrSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGlnGln 166  
||| :|||

Qy 528 CTGGCTACGGCTCTCTGTTTGGACACGTCGACGAGTTCTGGGAGCACCGCATGACTCG 587  
||| :|||  
Db 167 ValProTyrGlySerTrpTyrAspHisValLysAlaTrpTrpGluLysSerLysAsnSer 186  
||| :|||  
Qy 588 AACGTGCTTTTCTCAAGTATGAAGACATCATCGGACCTGCTGACGATCGTGGAGCAG 647  
||| :|||  
Db 187 ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgArgGluValValLys 206  
||| :|||  
Qy 648 CTGGCCAGATTCTCGGGGTGCTCTGTGACAAGCCCGAGCTGGAGCCCTGACGAGCAG 707  
||| :|||  
Db 207 LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleIleGlnHis 226  
||| :|||  
Qy 708 TGCAC-----CAGCTGGTGGACAG 728  
||| :|||  
Db 227 ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrMetMetProGluGlu 246  
||| :|||  
Qy 729 TGCTGCAACGCTGAGGCCCTGCC---GTGGCGCGGGAAGAGTTGGGCTGTGGAAGGAC 785  
||| :|||  
Db 247 MetMetAsnGlnLysValSerProPheMetArgLysGlyIleIleGlyAspTrpLysAsn 266  
||| :|||  
Qy 786 ATCTTCACCGTCTCCATGAATGAGAGTTTGTGACTTGTGTATATAACAGAGATGGGAAAG 845  
||| :|||  
Db 267 HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnGlnMetLysAsp 286  
||| :|||  
Qy 846 TGTGACCTCAGTTT 860  
||| :|||  
Db 287 CysThrValLysPhe 291  
||| :|||  
RESULT 14  
US-09-785-343-13  
; Sequence 13, Application US/09785343  
; Patent No. 6605455  
; GENERAL INFORMATION:  
; APPLICANT: The Board of Regents of the University of Oklahoma  
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF  
; FILE REFERENCE: 5827.003  
; CURRENT APPLICATION NUMBER: US/09/785,343  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: 09/150,133  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 60/072,994  
; PRIOR FILING DATE: 1998-01-29  
; PRIOR APPLICATION NUMBER: PCT/US99/16750  
; PRIOR FILING DATE: 1999-07-23  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)  
; SEQ ID NO 13  
; LENGTH: 295  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; FEATURE:  
; OTHER INFORMATION: estrogen sulfotransferase  
US-09-785-343-13  
Alignment Scores:  
Pred. No.: 2,28e-33 Length: 295  
Score: 430.00 Matches: 94  
Percent Similarity: 53.0% Conservative: 57  
Best Local Similarity: 33.0% Mismatches: 116  
Query Match: 9.8% Indels: 18  
Gaps: 3  
US-10-768-158-1 (1-2419) x US-09-785-343-13 (1-295)  
Qy 60 GAGTTCGAGCAAGTACTTCGAGTTCCATCGCTGGCGTCCGCCCTTCGCGCGG 119  
||| :|||  
Db 7 GluTyrGluValPheGlyGluPheArgGlyValLeuMetAspLysArgPheThrLys 26  
||| :|||  
Qy 120 AAGATGGAGGAGATCGCAACTTCCCGTGGCGGCCAGCGAGTGTGATCGTCACCTAC 179  
||| :|||  
Db 27 TyrTrpGluAspValGluMetPheLeuAlaArgProAspAspLeuValIleAlaThrTyr 46  
||| :|||

180	CCCAAGTCCGGCAGCAGCACTTCTGCTGAGGAGGTGGTCTACTTGGTGAGCCAGGGCGCTGAC	239
Qy		
47	ProLysSerGlyThrTrpIleSerGluValValTyrMetIleTyrLysGluGlyAsp	66
Db		
240	CCCGATGAGATCGGGCTTGATGAACATCGACGAGCAGCTCCCGGTCTCTGGAGTACCCACAG	299
Qy		
67	ValGluLysCysLysGluAspAlaIlePheAsnArgIleProTyrLeuGluCysArgAsn	86
Db		
300	CCG-----GGCTGGACATCATCAAGGAACATGACCTCTCTCCCGCTCATCAAG	347
Qy		
87	GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLysGluSerProArgIleValLys	106
Db		
348	AGCCACCTGCCCTACCGCTTCTGCGCTCTGACCTCCACAATGGAGACTCCCAAGTCCATC	407
Qy		
107	ThrHisLeuProLysValLeuProAlaSerPheTrpGluLysAsnCysLysMetIle	126
Db		
408	TATATGGCTCCCAACCCCAAGGATCTGGTGGTGCTTATTATCAGTTCCACCGCTCTCTG	467
Qy		
127	TyrLeuCysArgAsnAlaLysAspValalaValSerTyrTyrPheLeuLeuMetIle	146
Db		
468	CGGACCATAGACTACCGAGGCACCTTTCAAGAAATTTCTCCGGAGGTTTATGAATGATAG	527
Qy		
147	ThrSerTyrProAsnProLysSerPheSerGluPheValGluLysPheMetGlnGlyGln	166
Db		
528	CTGGGCTACGGCTCTGTGTTTGAGCAGCTGACGAGGAGTTCTGGGAGCACCAGCATGACTCG	587
Qy		
167	ValProTyrGlySerTrpTyrAspHisValLysAlaIleTrpGluLysSerLysAsnSer	186
Db		
588	AACGTGCTTTTTCACAGTATGAGACATGATCGGACCTGGTACGATCGTGGAGCAG	647
Qy		
187	ArgValLeuPheMetPheTyrGluAspMetLysGluAspIleArgArgGluValLys	206
Db		
648	CTGGCCAGATCTCTGGGGGTCTCTGTGACAAAGGCCAGCTGGGAAGCCCTGACGGAGCAC	707
Qy		
207	LeuIleGluPheLeuGluArgLysProSerAlaGluLeuValAspArgIleIleGlnHis	226
Db		
708	TGCCAC-----CAGTGTGTGGACACG	728
Qy		
227	ThrSerPheGlnGluMetLysAsnAsnProSerThrAsnTyrThrMetMetProGluGlu	246
Db		
729	TGCTGCAACGCTGAGGCCCTGCC-----CTGGCGCGGGAAGAGTTGGCTGTGGGAAGGAC	785
Qy		
247	MetMetAsnGlnLysValSerProPheMetArgLysGlyIleIleGlyAspTrpLysAsn	266
Db		
786	ATCTTACCCTCTCCATGAATCAGAAGTTTCACTTGGTGTATATAAACAAGAAGATGGAAAG	845
Qy		
267	HisPheProGluAlaLeuArgGluArgPheAspGluHisTyrLysGlnGlnMetLysAsp	286
Db		
846	TGTGACCTCAGCTTT	860
Qy		
287	CysThrValLysPhe	291
Db		

```

RESULT 15
US-10-411-976-13
; Sequence 13, Application US/10411976
; Patent No. 6713283
; GENERAL INFORMATION:
; APPLICANT: Moore, Kevin L.
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES
; FILE REFERENCE: 5864.027
; CURRENT APPLICATION NUMBER: US/10/411,976
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 09/785,343
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: estrogen sulfotransferase
US-10-411-976-13

```

Pred. No.:	2,28e-33	Length:	295
Score:	430.00	Matches:	54
Percent Similarity:	53.0%	Conservative:	97
Best Local Similarity:	33.0%	Mismatches:	116
Query Match:	9.8%	Indels:	18
DB:	2	Gaps:	3

US-10-768-158-1 (1-2419) x US-10-411-976-13 (1-295)

Qy	60	GAGTTCGAGAGCAAGTACTTCGAGTTCCATCGCGTGC	CGGCTCGCGCTTCCTTCGCGCGG	119
Db	7	GluTyrTyrGluValPheGlyGluPheArgGlyValLeu	MetAspLysArgPheThrLys	26
Qy	120	AAGATGGAGAGATCGCAACTTCCCGGTGGCGCCAC	GCGACGTCGTGGATCGTCACCTAC	179
Db	27	TyrTrpGluAspValGluMetPheLeuAlaArgProAsp	ProAspLeuValIleAlaThrTyr	46
Qy	180	CCCAAGTCCGACACGACTTCTCGACGAGAGTGTCT	ACTTGGTGAGCAGCGGCGCTGCAC	239
Db	47	ProLysSerGlyThrThrTrpIleSerGluValValTyr	MetIleTyrLysGluGlyAsp	666
Qy	240	CCCGATGATGATCGGCTTGATGAAATCGACGAGCAG	CTCCCGGTCTCGGAGTACCCACAC	299
Db	67	ValGluLysCysLysGluAspAlaIlePheAsnArgIle	ProTyrLeuGluCysArgAsn	86
Qy	300	CCG-----GGCTGGACATCATCAAGGAATGACCT	CTCTCCCGCTCATCAAG	347
Db	87	GluAspLeuIleAsnGlyIleLysGlnLeuLysGluLys	GluLysGluSerProArgIleValLys	106
Qy	348	AGCCACCTGCGCTACCGCTTTCGCCCTCTGACCTCA	CTCCACAATGGAGACTCCCAAGTCACTC	407
Db	107	ThrHisLeuProLysValLeuProAlaSerPheTrpGlu	LysAsnCysLysMetIle	126
Qy	408	TATATGGTCCCAACCCCAAGGATCTGTGTGTCTTAT	TATACAGTTCACCGCTCTCTG	467
Db	127	TyrLeuCysArgAsnAlaLysAspValAlaValSerTyr	TyrTyrPheLeuLeuMetIle	146
Qy	468	CGGACCATGAGCTACCGAGGACCTTTTCAAGAATTT	CTCCGGAGGTTTATGAATGATAAG	527
Db	147	ThrSerTyrProAsnProLysSerPheSerGluPheVal	GluLysPheMetGlnGlyGln	166
Qy	528	CTGGGCTACGGCTCGTGTGGTGGACACGTCGAGAGT	TCTTGGGACACCGCATGGACTCG	587
Db	167	ValProTyrGlySerTrpTyrAspHisValLysAlaTrp	TrpGluLysSerLysAsnSer	186
Qy	588	AACGTGCTTTTCTCAAGTATGAAGACATGCATCGG	ACCTGGTGACGATGGTGAGCAG	647
Db	187	ArgValLeuPheMetPheTyrGluAspMetLysGluAsp	IleArgArgGluValValLys	206
Qy	648	CTGGCCAGATCTCTGGGGGTCTCCTGTGACAAAGG	CCCAAGTCCGACGCGTACGAGCAG	707
Db	207	LeuIleGluPheLeuGluArgLysProSerAlaGluLeu	ValAspArgIleIleGlnHis	226
Qy	708	TGCCAC-----	-----CAGCTGGTGACACG	728
Db	227	ThrSerPheGlnGluMetLysAsnAsnProSerThrAsn	TyrThrMetMetProGluGlu	246
Qy	729	TGCTGCAACGCTAGAGCCCTGCC- -GTGGGCGGG	GAAGAGTGGCTGTGGGAAGAC	785
Db	247	MetMetAsnGlnLysValSerProPheMetArgLysGly	IleIleGlyAspTrpLysAsn	266
Qy	786	ATCTTTCACCGTCTCCATGAATGAGAAGTTTCACT	TGGTGTATATAACAGAAGATGGGAAG	845
Db	267	HisPheProGluAlaLeuArgGluArgPheAspGluHis	TyrLysGlnGlnMetLysAsp	286
Qy	846	TGTGACCTCAGTTT	860	
Db	287	CysThrValLysPhe	291	

RESULT 16  
US-10-201-525-13  
; Sequence 13, Application US/10201525  
; Patent No. 7029890

RESULT 16  
US-10-201-525-13  
; Sequence 13, Application US/10201525  
; Patent No. 7029890





Dbb 81 AlaGlnThrLeuAspArgHisAlaPheLeuLeuLeuLysPheProHisLysGluLysPro 100  
Qy 303 GGCCTGGACATCATCAAGGAACCTACCTCTCCCGCCCTCATCAAGAGCACCTGCCCTAC 362  
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuLeuLysThrHisLeuProSer 120  
Qy 363 CGCTTTCTGCCCTCTGACCTCCACATCGAGAGCTCAAGGTCATCTATATATGGCTCGCAAC 422  
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyValAlaArgAsn 140  
Qy 423 CCCAAGGATCTGGTGTCTTATATCATGTTCCACCGCTCTCTCGGACCATGAGCTAC 482  
Db 141 ProllysAspCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAsp 160  
Qy 483 CGAGGCACCTTTCAAGATCTCGCGAGGTTTATGAATGATAAGCTGGCTACGGCTCC 542  
Db 161 ProGlnAsnLeuGluGluPheTyTrpLysPheMetSerGlyLysValValGlySer 180  
Qy 543 TGGTTTGAGCAGTGCAGGAGTCTTGGAGCACCGCATGGACTCGAAGCTGCTTTTCTC 602  
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200  
Qy 603 AAGTATGAACATCATCGGAGCTGTGAGCATGGTGGAGCAGCTGGCCAGATCTCTG 662  
Db 201 PheTyrGluAspIleLysLysAspProllysArgGluIleGluLysIleLeuLysPheLeu 220  
Qy 663 GGGGTGTCTGTGACAGGCCAGCTGGAAGCCCTGACGGAGCACTGCCACAGCTGTGT 722  
Db 221 GluLysAspIleSerGluGluIleLeuAsnLysIleIleTyrHisThrSerPheAspVal 240  
Qy 723 -----GACCAGTGTGCAAGCTGAGCGCCCTGCGCGTGGGC----- 758  
Db 241 MetLysGlnAsnProMetThrAsnTyrThrThrLeuProThrSerIleMetAspHisSer 260  
Qy 759 -----CGGGAGAGATTGGGCTGGAGGACATCTTCCACCGTCTCC 800  
Db 261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAla 280  
Qy 801 ATGAATGAGAAGTTGACTTGGTGTATATAAAGAGAGATGGGAAAGTGTGACCTCACGTTT 860  
Db 281 GlnAsnGluGluPheAspLysAspTyrGlnLysLysMetAlaGlySerThrLeuThrPhe 300

RESULT 18

US-10-199-334-5  
; Sequence 5, Application US/10199334  
; Patent No. 6905855  
; GENERAL INFORMATION:  
; APPLICANT: WOODAGE, Trevor et al.  
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING  
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN  
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL000669DIV-3  
; CURRENT APPLICATION NUMBER: US/10/199,334  
; CURRENT FILING DATE: 2002-07-22  
; PRIOR APPLICATION NUMBER: 09/609,816  
; PRIOR FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: 60/192,408  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: 60/212,725  
; PRIOR FILING DATE: 2000-06-20  
; PRIOR APPLICATION NUMBER: 09/609,816  
; PRIOR FILING DATE: 2000-07-03  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 304  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-199-334-5

Alignment Scores: 1.43e-32 Length: 304  
Pred. No.: 422.00 Matches: 97  
Score:

Percent Similarity: 52.0% Conservative: 59  
Best Local Similarity: 32.3% Mismatches: 124  
Query Match: 9.6% Indels: 20  
DB: 2 Gaps: 5  
US-10-768-158-1 (1-2419) x US-10-199-334-5 (1-304)  
Qy 21 ATGCGGAGACGAGCGGAGACCCAGCACCCAGCACCCCGGGGAGTTCGAG---AGCAAGTAC 77  
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20  
Qy 78 TTCAGATTCCATGCGGTGCGGCTCCGCCCTTCTGCGCGGGAAGATGAGGAGATCGCC 137  
Db 21 MetGluValAspGlyValProThrLeuLeuSerLysGluTrpTrpGluLysValCys 40  
Qy 138 AACTTCCCGGTGCGGCGGAGCGAGCTGTGGATCTGACCTACCCCAAGTCCGGCACCAGC 197  
Db 41 AsnPheGlnAlaLysProAspAspLeuIleAlaThrTyrProllysSerGlyThrThr 60  
Qy 198 TTGCTGAGGAGTGTCTACTTGTGAGCGGCGCTGACCCCGATGAGATCGGCTTG 257  
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80  
Qy 258 ATGAACATCGACGAGAGCTCCCGTCTCGAG-----TACCCA-----CAGCCG 302  
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysGluLysPro 100  
Qy 303 GGCCTGGACATCATCAAGGAACCTGACCTCTCCCGCCCTCATCAAGAGCACCTGCCCTAC 362  
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuLysIleIleTyrHisThrSer 120  
Qy 363 CGCTTTCTGCCCTCTGACCTCCACCAATGGAGACTCCCAAGGTCATCTATATGCGTCGCAAC 422  
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyValAlaArgAsn 140  
Qy 423 CCCAAGGATCTGGTGTCTTATATCATGTTCCACCGCTCTCTCGGACCATGAGCTAC 482  
Db 141 ProllysAspCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAsp 160  
Qy 483 CGAGGCACCTTTCAAGAAATTTGCGCGAGGTTTATGAATGATAAGCTGGGCTACGGCTCC 542  
Db 161 ProGlnAsnLeuGluGluPheTyTrpLysPheMetSerGlyLysValValGlySer 180  
Qy 543 TGGTTTGAGCAGTGCAGGAGTCTTGGAGCACCGCATGGACTCGAAGCTGCTTTTCTC 602  
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200  
Qy 603 AAGTATGAACATCATCGGAGCTGTGAGCATGGTGGAGCAGCTGGCCAGATCTCTG 662  
Db 201 PheTyrGluAspIleLysLysAspProllysArgGluIleGluLysIleLeuLysPheLeu 220  
Qy 663 GGGGTGTCTGTGACAGGCCAGCTGGAAGCCCTGACGGAGCACTGCCACAGCTGTGT 722  
Db 221 GluLysAspLysSerGluGluIleLeuAsnLysIleIleTyrHisThrSerPheAspVal 240  
Qy 723 -----GACCAGTGTGCAAGCTGAGCGCCCTGCGCGTGGGC----- 758  
Db 241 MetLysGlnAsnProMetThrAsnTyrThrThrLeuProThrSerIleMetAspHisSer 260  
Qy 759 -----CGGGAGAGATTGGGCTGGAGGACATCTTCCACCGTCTCC 800  
Db 261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAla 280  
Qy 801 ATGAATGAGAAGTTTGAATGTTGGTGTATATAAAGAGAGATGGGAAAGTGTGACCTCACGTTT 860  
Db 281 GlnAsnGluGluPheAspLysAspTyrGlnLysLysMetAlaGlySerThrLeuThrPhe 300

RESULT 19

US-10-199-329-5  
; Sequence 5, Application US/10199329  
; Patent No. 6953681  
; GENERAL INFORMATION:  
; APPLICANT: WOODAGE, Trevor et al.

```
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-199-329-5

Alignment Scores:
Pred. No.: 1,43e-32 Length: 304
Score: 422.00 Matches: 97
Percent Similarity: 52.0% Conservative: 59
Best Local Similarity: 32.3% Mismatches: 124
Query Match: 9.6% Indels: 20
DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x US-10-199-329-5 (1-304)
QY 21 ATGCGGAGAGCGAGCGGAGACCCAGACCCCGGGGGAGTTCGAG---AGCAAGTAC 77
DB 1 MetAlaLysIleGluLeuAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
QY 78 TTCGAGTTCCTCCATCAAGGAACTGACCTCTCCCGCCCTCATCAAGAGCCACTGCCCTAC 137
DB 21 MetGluValAspGlyValProThrLeuLeuSerLysGluTrpTrpGluLysValCys 40
QY 138 AACTTCCCGTGGCGGCGGAGCGAGTGTGGTGTGATGCTACCTACCCCAAGTCCGGACACG 197
DB 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrProLysSerGlyThrThr 60
QY 198 TTGTCGAGGAGGTGCTTACTTGTGTGAGCGAGCGGCGTGTGAGCGAGTGTGAGTGTG 257
DB 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
QY 258 ATGAACATCGACGAGCACTCCCGTCTCTGAG---TACCCA-----CAGCGG 302
DB 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
QY 303 GGCCTGGACATCATCAAGGAACTGACCTCTCCCGCCCTCATCAAGAGCCACTGCCCTAC 362
DB 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
QY 363 CGCTTTCGCTCTGACCTCCACATCGAGACTCCAGGTCATCATATATGCTGCGAAC 422
DB 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValIleValAlaArgAsn 140
QY 423 CCCAAGGATCGGTGCTTATTATTCAGTTCCACCGCTCTCTCGGACCATGAGTAC 482
DB 141 ProLysAspCysLeuValSerIleThrHisPheHisArgMetAlaSerPheMetProAsp 160
QY 483 CGAGGCACCTTTCAGAAATTCGCGGAGGTTTATGAATGATTAAGTGGGTTCGCTCC 542
DB 161 ProGlnAsnLeuGluGluPheTrpGluLysPheMetSerGlyLysValValGlyCys 180
QY 543 TGGTTTGGACGCTCAGGATTCGAGGACCGCATGGACTCGAACTGCTTTTCTC 602
DB 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTrpLeu 200
QY 603 AAGTATGAACATGCATCGGACCTGCTGACGATGGTGGAGCAGCTGGCCAGATTCCTG 662
```

```
DB 201 PheTyrGluAspIleLysLysAspProLysArgGluIleGluLysIleLeuLysPheLeu 220
QY 663 GGGGTGTCCTGTGACAGAGCCAGCTGGAAGCCCTGACGAGGACTCCACGACTCGTG 722
DB 221 GluLysAspIleSerGluGluIleLeuAsnLysIleIleTyrHisThrSerPheAspVal 240
QY 723 -----GACCAGTGTCTCAACGCTGAGGCGCTGCCCGTGGC----- 758
DB 241 MetLysGlnAsnProMetThrAsnTyrThrThrLeuProThrSerIleMetAspHisSer 260
QY 759 -----CGGGAAGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCC 800
DB 261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAla 280
QY 801 ATGAATCAGAAAGTTGACTGTGGTGTATAAACAGAAATGGAAAGTGTGACCTCAGCTTT 860
DB 281 GlnAsnGluGluPheAspLysAspTyrGlnLysLysMetAlaGlySerThrLeuThrPhe 300
```

## RESULT 20

```
US-10-199-330-5
; Sequence 5, Application US/10199330
; Patent No. 6967094
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-199-330-5
```

## Alignment Scores:

```
Pred. No.: 1,43e-32 Length: 304
Score: 422.00 Matches: 97
Percent Similarity: 52.0% Conservative: 59
Best Local Similarity: 32.3% Mismatches: 124
Query Match: 9.6% Indels: 20
DB: 2 Gaps: 5
```

US-10-768-158-1 (1-2419) x US-10-199-330-5 (1-304)

```
QY 21 ATGCGGAGAGCGAGCGGAGACCCAGACCCCGGGGGAGTTCGAG---AGCAAGTAC 77
DB 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
QY 78 TTCGAGTTCCTCCATCAAGGAACTGACCTCTCCCGCCCTCATCAAGAGCCACTGCC 137
DB 21 MetGluValAspGlyValProThrLeuLeuSerLysGluTrpTrpGluLysValCys 40
QY 138 AACTTCCCGTGGCGGCGGAGCGAGTGTGGTGTGATGCTACCTACCCCAAGTCCGGACACG 197
DB 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrProLysSerGlyThrThr 60
QY 198 TTGTCGAGGAGGTGCTTACTTGTGTGAGCGGCGTGTGAGCGAGTGTGAGTGTG 257
DB 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
```

```
QY 258 ATGAACATCGACGACGACTCCCGCTCTGGAG-----TACCCA-----CAGCCG 302
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 303 GGCCTGGACATCAAGGAATGACCTCCCGCTCATCAAGAGCCACTGCCCTAC 362
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 101 AspLeuGluPheValLeuGluMetSerProGlnLeuIleLysThrHisLeuProSer 120
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 363 CGCTTTCTGCTCTGACCTCCCAATGGAGACTCCAAAGTCACTATATATGGCTCGCAAC 422
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 423 CCAAGGATCTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGGACCATGAGTAC 482
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 141 ProLysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSerPheMetProAsp 160
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 483 CGAGGCACCTTTCAAGAACTTCGCGGAGTTTATGAATGATAGTACGGCTACGGCTCC 542
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValIleGlyLys 180
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 543 TGGTTTGAGCACGTGCGAGAGTTCTGGGAGCACCGCATGGACTCGAAAGTGTCTTTTCTC 602
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 181 TrpPheAspHisValLysGlyTrpTrpAlaLysAspMetHisArgIleLeuTyrLeu 200
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 603 AAGTATGAACATGATCGGACCTGTGTGACGATGGTGGAGCAGCTGGCCAGATTCCTG 662
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 201 PheTyrGluAspIleLysLysAspProLysArgGluIleGluLysIleLeuLysPheLeu 220
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 663 GGGGTGCTCTGACAAGCCAGCTGGAAGCCCTGACGGAGCAGTGCACCAGCTGGTG 722
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 221 GluLysAspIleSerGluGluIleLeuAsnLysIleIleTyrHisThrSerPheAspVal 240
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 723 -----GACCAGTGTGCAACGCTGAGGCCCTGCCCGTGGG----- 758
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 241 MetLysGlnAsnProMetThrAsnTyrThrThrLeuProThrSerIleMetAspHisSer 260
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 759 -----CGGGGAAGATTGGGTGTGGAGGACATCTTACCGTCTCC 800
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 261 IleSerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAla 280
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 801 ATGAATGAGAAGTTGACTTGGTGATATAACAGAGATGGGAAGTGAGACCTCACGTTT 860
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 281 GlnAsnGluGluPheAspLysAspTyrGlnLysLysMetAlaGlySerThrLeuThrPhe 300
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
```

## RESULT 21

```
US-09-609-816-8
; Sequence 8, Application US/09609816
; Patent No. 6436884
; GENERAL INFORMATION:
; APPLICANT: Woodage, Trevor
; APPLICANT: Wei, Minh Hui
; APPLICANT: Kodira, Chinappa
; APPLICANT: Beasley, Ellen
; APPLICANT: DiFrancesco, Valentina
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO00669PCT
; CURRENT APPLICATION NUMBER: US/09/609,816
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: HUMAN
; APPLICANT: Woodage, Trevor
```

```
Alignment Scores:
Pred. No.: 2,42e-32 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.6% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: 2 Gaps: 4
```

US-10-768-158-1 (1-2419) x US-09-609-816-8 (1-283)

```
QY 81 GAGTTCATCGCTGCGCTGCCCGCTTTCGCCGCGGAAGATGGAGGAGATGCCCAAC 140
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 141 TTCCCGGTGCGCGCCAGCGACGTGTGTGATCGTACCTACCCCAAGTCCGGCCAGCAGCTTG 200
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 21 PheGlnAlaLysProAspAspLeuLeuAlaThrTyrAlaLysAlaGlyThrTrp 40
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 201 CTGACGAGGTGTCTACTTGGTGAGCCAGCGGCTGACCCCGATGAGATCGGCTTGATG 260
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 261 AACATCCAGCAGCAGCTCCCGCTCTCGAGTAC-----CCACAGCCG-----GGC 305
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 306 CTGACATCATCAAGGAAGTACCTCTCCCGCTCAAGAGCCACTGCCCTACCGC 365
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 366 TTCTGCGCTCTGACCTCCCAATGGAGACTCCAAAGTCTATATATGGTTCGCAACCCC 425
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 101 MetLeuProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAsnAla 120
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 426 AAGATCTGTGTGTCTTATTATCAGTTCCACCGCTCTCTGCGGACCATGAGTACCGA 485
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 486 GGCACTTTCAGAAATCTGCCGGAGTTTATGAATGATAGTGGGTGGGTCTCGCTCCG 545
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 141 GlyThrLeuGlyGluTyrIleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 160
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 546 TTGAGCAGCTGCGAGGATTTCTGGGAGCACCGCATGGACTCGAACGTCTTTTCTCAAG 605
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 161 TyrAspHisValLysGlyTrpTrpAspValLysAspGlnHisArgIleLeuTyrLeuPhe 180
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 606 TATGAACATGATCGGACCTCGGACCTGTCACGATGGTGGAGCAGCTGGCCAGATTCTTGGGG 665
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 200
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 666 GTGTCTGTGACAAGGCCAGCTCGAAGCCCTGCGGAGCAGCTCCACCACTGGTG-- 722
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 201 LysAspIleSerGluGluValLeuAsnLysIleIleTyrHisThrSerPheAspValMet 220
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 723 -----GACCAGTGTCTGCAACGCTGAGGCCCTGCCCGCTGCC----- 752
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 221 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 753 -----GTGGCCCGGGGAAGATTTGGGTGTGGAGGAGCATCTTCACCGTCTCCATG 803
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 804 AATGAGAAGTTTGGTGTATATAACAGAAAGATGGAAAGTGTGACCTCACGTTT 860
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
QY 261 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 279
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
```

## RESULT 22

```
US-09-609-816-9
; Sequence 9, Application US/09609816
; Patent No. 6436884
; GENERAL INFORMATION:
; APPLICANT: Woodage, Trevor
```



```
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
Qy 366 TTTCTGCCCTCTGACCTCCACAATGGAGACATCCCAAGGTCACTATATATGCTCGCAACCCC 425
Db 101 MetLeuProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAsnAla 120
Qy 426 AAGATCTGGTGGTCTTATATATCAGTTCCACCGCTCTCTCGGACCATCAGTACCGA 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
Qy 486 GGCACCTTTCAAGAAATCTCGCGAGGTTTATGAATGATGAAGCTAGCTCGGCTCCCTGG 545
Db 141 GlyThrLeuGlyGluTyrIleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 160
Qy 546 TTTGAGCAGTCGACGAGTCTTGGAGCACCACCGCATGGACTCGAAGCTGCTTTTCTCAAG 605
Db 161 TyrAspHisValLysGlyTrpAspValLysAspGlnHisArgIleLeuTyrLeuPhe 180
Qy 606 TATGAACATGTCATCGGACCTGGTGCAGATGGTGGAGCAGCTGGCCAGATTCTCTGGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 200
Qy 666 GTGTCTGTGACAAAGCCACCTGGAAGCCCTGACGGAGCACTGCCACCAGCTGTG-- 722
Db 201 LysAspLysSerGluGluValLeuAsnLysIleIleTyrHisThrSerPheAspValMet 220
Qy 723 -----GACCAGTGTGCAACGCTGAGCCCTCGCC----- 752
Db 221 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240
Qy 753 -----GTGGCCGGGGAAGAGTTGGGCTGTGGAAGAGCATCTTCACCGTCTCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
Qy 804 AATGAGAAGTTTGACTTGGTGTATAAACAGAGAGATGGAAAGTGTGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 279

RESULT 24
US-10-199-334-9
; Sequence 9, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-9

Alignment Scores:
Pred. No.: 2,42e-32 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.8% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: 2 Gaps: 4
```

```
US-10-768-158-1 (1-2419) x US-10-199-334-9 (1-283)
Qy 81 GAGTTTCATCGCTGCGCTGCCCGCTTCTGCCGCGGAAGATGGAGAGATCGCCCAAC 140
Db 1 GluValAsnGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
Qy 141 TTCCCGTGGCGGCCAGCGAGCTGTGATCGTACCTACCCCAAGTCCGCGACCAAGCTTG 200
Db 21 PheGlnAlaLysProAspAspLeuLeuIleAlaThrTyrAlaLysAlaGlyThrThrTrp 40
Qy 201 CTGCAGAGGTGGTCTACTTGGTGAGCCGCGCTGACCCCGATGAGATCGGCTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
Qy 261 AACATCGACGACGAGCTCCCGTCTCTGAGTAC-----CCACAGCCG-----GGC 305
Db 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
Qy 306 CTGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACCTGCCCTACCGC 365
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
Qy 366 TTTCTGCCCTCTGACCTCCACAATGGAGACTCCCAAGGTCACTATATATGGCTCGCAACCCC 425
Db 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAsnAla 120
Qy 426 AAGCATCTGGTGGTGTCTTATTATCAGTTCCACCGCTCTCTGCGGACCATCAGCTACCGA 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
Qy 486 GGCACTTTCAAGAAATCTCGCGAGGTTTATGAATGATGAAGTGGGCTGCGCTCTCTGGG 545
Db 141 GlyThrLeuGlyGluTyrIleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 160
Qy 546 TTTGAGCAGTCGACGAGTCTTGGAGCACCACCGCATGGACTCGAAGCTGCTTTTCTCAAG 605
Db 161 TyrAspHisValLysGlyTrpTrpAspValLysAspGlnHisArgIleLeuTyrLeuPhe 180
Qy 606 TATGAACATGTCATCGGACCTGGTGCAGTGTGTGAGCAGCTGGCCAGATTTCTTGGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 200
Qy 666 GTGTCTGTGACAAAGCCCGAGCTGGAAGCCCTGACGGAGCACTGCCACCAGCTGTG--- 722
Db 201 LysAspLysSerGluGluValLeuAsnLysIleIleTyrHisThrSerPheAspValMet 220
Qy 723 -----GACCAGTGTGCAACGCTGAGCCCTCGCC----- 752
Db 221 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240
Qy 753 -----GTGGCCGGGGAAGAGTTGGGCTGTGGAAGAGCATCTTCACCGTCTCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
Qy 804 AATGAGAAGTTTGACTTGGTGTATAAACAGAGAGATGGAAAGTGTGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 279

RESULT 25
US-10-199-329-8
; Sequence 8, Application US/10199329
; Patent No. 6953681
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
```

```
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-8

Alignment Scores:
Pred. No.:      2,42e-32      Length:      283
Score:          419.50      Matches:      91
Percent Similarity: 53.0%      Conservative: 57
Best Local Similarity: 32.6%      Mismatches: 112
Query Match:    9.5%         Indels:      19
DB:             2           Gaps:         4

US-10-768-158-1 (1-2419) x US-10-199-329-8 (1-283)
Qy 81 GAGTTCATGCGTGGCGCTCCGCCCTTCTGCGCGGAAGATGGAGGATCGCCAAAC 140
Db 1 GluValAenGlyIleLeuMetSerLysLeuMetSerAspAenTrpAspLysIleTrpAsn 20
Qy 141 TTCCCGGTGCGCGCGAGCTGGGATCGTCACTACCCCAAGTCCGGCCAGCAGCTTG 200
Db 21 PheGlnAlaLysProAspAspLeuLeuAlaThrTyrAlaLysAlaGlyThrTrp 40
Qy 201 CTGCAGGAGGTGCTACTTGGTGAGCCAGGCGCTGACCCGATGAGATCGGCTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
Qy 261 AACATCGACGACGAGCTCCCGTCTCGAGTAC-----CCACAGCCG-----GGC 305
Db 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
Qy 306 CTGCACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACTGCCCTACCGC 365
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
Qy 366 TTTCGCGCTCTGACCTCCACATGGAGCTCCCAAGTCTATATATGGCTCGCAACCCC 425
Db 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAla 120
Qy 426 AAGGATCTGGTGTGCTTATATCAGTTCCACCGCTCTCTCGGACCATGAGTACCGA 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140

US-10-768-158-1 (1-2419) x US-10-199-329-9 (1-283)
Qy 81 GAGTTCATGCGTGGCGCTCCGCCCTTCTGCGCGGAAGATGGAGGATCGCCAAAC 140
Db 1 GluValAenGlyIleLeuMetSerLysLeuMetSerAspAenTrpAspLysIleTrpAsn 20
Qy 141 TTCCCGGTGCGCGCGAGCTGGGATCGTCACTACCCCAAGTCCGGCCAGCAGCTTG 200
Db 21 PheGlnAlaLysProAspAspLeuLeuAlaThrTyrAlaLysAlaGlyThrTrp 40
Qy 201 CTGCAGGAGGTGCTACTTGGTGAGCCAGGCGCTGACCCGATGAGATCGGCTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
Qy 261 AACATCGACGACGAGCTCCCGTCTCGAGTAC-----CCACAGCCG-----GGC 305
Db 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
Qy 306 CTGCACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACTGCCCTACCGC 365
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
Qy 366 TTTCGCGCTCTGACCTCCACATGGAGCTCCCAAGTCTATATATGGCTCGCAACCCC 425
Db 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAla 120
Qy 426 AAGGATCTGGTGTGCTTATATCAGTTCCACCGCTCTCTCGGACCATGAGTACCGA 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140

US-10-199-329-9
; Sequence 9, Application US/10199329
; Patent NO. 6953681
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-9

Alignment Scores:
Pred. No.:      2,42e-32      Length:      283
Score:          419.50      Matches:      91
Percent Similarity: 53.0%      Conservative: 57
Best Local Similarity: 32.6%      Mismatches: 112
Query Match:    9.5%         Indels:      19
DB:             2           Gaps:         4

US-10-199-329-9 (1-2419) x US-10-199-329-8 (1-283)
Qy 81 GAGTTCATGCGTGGCGCTCCGCCCTTCTGCGCGGAAGATGGAGGATCGCCAAAC 140
Db 1 GluValAenGlyIleLeuMetSerLysLeuMetSerAspAenTrpAspLysIleTrpAsn 20
Qy 141 TTCCCGGTGCGCGCGAGCTGGGATCGTCACTACCCCAAGTCCGGCCAGCAGCTTG 200
Db 21 PheGlnAlaLysProAspAspLeuLeuAlaThrTyrAlaLysAlaGlyThrTrp 40
Qy 201 CTGCAGGAGGTGCTACTTGGTGAGCCAGGCGCTGACCCGATGAGATCGGCTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
Qy 261 AACATCGACGACGAGCTCCCGTCTCGAGTAC-----CCACAGCCG-----GGC 305
Db 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
Qy 306 CTGCACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACTGCCCTACCGC 365
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
Qy 366 TTTCGCGCTCTGACCTCCACATGGAGCTCCCAAGTCTATATATGGCTCGCAACCCC 425
Db 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAla 120
Qy 426 AAGGATCTGGTGTGCTTATATCAGTTCCACCGCTCTCTCGGACCATGAGTACCGA 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140

US-10-199-329-9 (1-2419) x US-10-199-329-9 (1-283)
Qy 81 GAGTTCATGCGTGGCGCTCCGCCCTTCTGCGCGGAAGATGGAGGATCGCCAAAC 140
Db 1 GluValAenGlyIleLeuMetSerLysLeuMetSerAspAenTrpAspLysIleTrpAsn 20
Qy 141 TTCCCGGTGCGCGCGAGCTGGGATCGTCACTACCCCAAGTCCGGCCAGCAGCTTG 200
Db 21 PheGlnAlaLysProAspAspLeuLeuAlaThrTyrAlaLysAlaGlyThrTrp 40
Qy 201 CTGCAGGAGGTGCTACTTGGTGAGCCAGGCGCTGACCCGATGAGATCGGCTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
Qy 261 AACATCGACGACGAGCTCCCGTCTCGAGTAC-----CCACAGCCG-----GGC 305
Db 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
Qy 306 CTGCACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCCACTGCCCTACCGC 365
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
Qy 366 TTTCGCGCTCTGACCTCCACATGGAGCTCCCAAGTCTATATATGGCTCGCAACCCC 425
Db 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAla 120
Qy 426 AAGGATCTGGTGTGCTTATATCAGTTCCACCGCTCTCTCGGACCATGAGTACCGA 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
```

```
QY 486 GGCACCTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCCTGG 545
Db 141 GlyThrLeuGlyGluTyrlleGlulnPhelYsAlaGlyLysValLeuTrpGlySerTrp 160
QY 546 TTTGAGCAGTCGACGAGTTCTGGGAGCACCGCATGGACTCGAAGTCTCTTTTCTCAAG 605
Db 161 TyrAspHisValYsGlyTrpTrpAspValYsAspGlnHisArgIleuTyrlleuPhe 180
QY 606 TATGAAGACATGCATCGGACCTGGTGACGATGGTGAGCAGCTGGCCAGATTCCTGGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 200
QY 666 GTGCTCTGTACAGGCCAGCTGGAAGCCCTGACGGAGCAGCTGCCACCACTGGTG--- 722
Db 201 LysAspIleSerGluGluValLeuAsnLysIleIleTyrlleHisThrSerPheAspValMet 220
QY 723 -----GACCAGTGTGCAACGCTGAGGCCCTGCC----- 752
Db 221 LysGluAsnProMetAlaAsnTyrlleThrLeuProSerSerIleMetAspHisSerIle 240
QY 753 -----GTGGCGCGGGAAGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrlleThrValAlaGln 260
QY 804 AATGAGAAGTTTGACTTGGTGATATAACAGAGATGGGAAGTGTGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrlleArgLysMetAlaGlySerAsnIleThrPhe 279

RESULT 27
US-10-199-330-8
; Sequence 8, Application US/10199330
; Patent No. 6967094
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-8

Alignment Scores:
Pred. No.: 2, 42e-32 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.6% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: 2 Gaps: 4

US-10-768-158-1 (1-241) x US-10-199-330-8 (1-283)
QY 81 GAGTTCCATGGCGTGGCTGCGCCCTTCTGCGCGGGAAGATGAGGAGATGCCCAAC 140
Db 1 GluValAsnGlyIleLeuMetSerLysLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
QY 141 TTCGCGTGGCGGCCAGGACGTCGTGGATCGTCACCTACCCCAAGTCGGGACCACTGG 200
Db 21 PheGlnAlaLysProAspAspLeuIleAlaThrTyrlleAlaLysAlaGlyThrThrTrp 40
```

```
QY 201 CTGACGAGGTGGTCTACTTGGTGAGCGAGGGCCCTGACCCCGATGAGATCGGCTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
QY 261 AACATCGACGAGCAGCTCCCGTCTCTGAGTAC-----CCACAGCCG-----GGC 305
Db 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
QY 306 CTGACATCATCAAGAACTGACCTCTCCCGCCTCATCAAGAGCCACCTGCCCTACCGC 365
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValHis 100
QY 366 TTTTGGCCTCTGACCTCCACATGAGACTCCCAAGTCTCATATGATGCTGCGCAACCCC 425
Db 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrlleValAlaArgAsnAla 120
QY 426 AAGCATCTGTGGTGTCTTATTATCAGTCTCCACCGCTCTCTGCGGACCATGAGCTACCGA 485
Db 121 LysAspCysLeuValSerTyrlleThrPheSerArgMetAsnLysMetLeuProAspPro 140
QY 486 GGCACCTTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTGGGCTCCTGG 545
Db 141 GlyThrLeuGlyGluTyrlleGluGlnPheLysAlaGlyLysValLeuTrpGlySerTrp 160
QY 546 TTTGAGCAGTCGACGAGTTCTGGGAGCACCGCATGGACTCGAAGCTGCTTTTCTCAAG 605
Db 161 TyrAspHisValLysGlyTrpTrpAspValLysAspGlnHisArgIleLeuTyrllePhe 180
QY 606 TATGAAGACATGCATCGGACCTGGTGACGATGGTGAGCAGCTGGCCAGATTCCTGGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 200
QY 666 GTGCTCTGTGACAAAGCCAGCTGGAAGCCCTGAGGAGCAGCTGCCACCACTGGTG--- 722
Db 201 LysAspIleSerGluGluValLeuAsnLysIleIleTyrlleHisThrSerPheAspValMet 220
QY 723 -----GACCAGTGTGCAACGCTGAGGCCCTGCC----- 752
Db 221 LysGluAsnProMetAlaAsnTyrlleThrLeuProSerSerIleMetAspHisSerIle 240
QY 753 -----GTGGCGCGGGAAGAGTTGGGCTGTGGAAGGACATCTTCACCGTCTCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrlleThrValAlaGln 260
QY 804 AATGAGAAGTTTGACTTGGTGATATAACAGAGATGGGAAGTGTGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrlleArgLysMetAlaGlySerAsnIleThrPhe 279
```

```
RESULT 28
US-10-199-330-9
; Sequence 9, Application US/10199330
; Patent No. 6967094
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 283
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-9

Alignment Scores:
Pred. No.: 2,42e-32 Length: 283
Score: 419.50 Matches: 91
Percent Similarity: 53.0% Conservative: 57
Best Local Similarity: 32.6% Mismatches: 112
Query Match: 9.5% Indels: 19
DB: 2 Gaps: 4

US-10-768-158-1 (1-2419) x US-10-199-330-9 (1-283)
Qy 81 GAGTTCATGCGTGGCGCTCGCCCTTTCGCCGGGAAGATGGAGGATGCCCAAC 140
Db 1 GluValAenGlyIleLeuMetSerLysLeuMetSerAspAsnTrpAspLysIleTrpAsn 20
Qy 141 TTCGGGTGGCGCCGAGCGTGGATCGTACCTACCCCAAGTCCGGCAGCAGCTTG 200
Db 21 PheGlnAlaLysProAspAspLeuLeuIleAlaThrYrAlaLysAlaGlyThrTrp 40
Qy 201 CTGAGGAGGTGCTACTTGGTGAGCCAGCGCGTACCCCGATGAGATCGGCTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
Qy 261 AACATCGACGAGCTCCCGGTCTGGAGTAC-----CCACAGCG-----GGC 305
Db 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProSerProLeuAsnSerGly 80
Qy 306 CTGACATCATCAAGAACTGACCTCCCGCTCATCAAGAGCCACCTGCCCTACCGC 365
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValGln 100
Qy 366 TTTCTGCCCTCTGACCTCCCAAGTCCCAAGTCTATATGATGAGTGGTCCGAACCCC 425
Db 101 MetLeuProProSerPheTrpLysGluAsnSerLysIleIleTyrValAlaArgAsnAla 120
Qy 426 AAGGATCGTGGTGTCTTATTATCAGTTCCACCGCTCTCTCGCGAGCATGAGCTACCGA 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuLeuProAspPro 140
Qy 486 GGCACCTTTCAAGAAATTTCTGCCGGAGGTTTATGAATGATGAAGCTGGCTCTGG 545
Db 141 GlyThrLeuGlyGluTyrIleGluThrPheLysAlaGlyLysValLeuTrpGlySerTrp 160
Qy 546 TTTGAGACGTCGAGAGTCTCTGGAGACCGCATGAGTGGTCCAGCTCTTTCTCAAG 605
Db 161 TyrAspHisValLysGlyTrpAspValLysAspGlnHisArgIleLeuTyrLeuPhe 180
Qy 606 TATGAAGACATGCAATCGGAGCTGGTACGATGGTGGAGAGCTGGCCAGATTCTCTGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleAlaLysPheLeuGlu 200
Qy 666 GTGCTCTGTGACAGGCCACGTGAACCTGACGGAGCACTGCCACCACTGGTG-- 722
Db 201 LysAspIleSerGluGluValLeuAsnLysIleIleTyrHisThrSerPheAspValMet 220
Qy 723 -----GACAGTGTGCAAGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCG 752
Db 221 LysGluAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240
Qy 753 -----GTGGCGCGGGAAGATTGGCTGGTGGAGGAGCATTTTCCCGTCTCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
Qy 804 AATGAGAAGTTTGAATCTGGTGTATAAACAGAAAGTGGGAAAGTGGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrArgArgLysMetAlaGlySerAsnIleThrPhe 279

RESULT 29
US-09-609-816-10
; Sequence 10, Application US/09609816
```



```
Qy 606 TATGAGACATGCTGGGACCTGGTGCAGATGTTGGAGCAGCTGGCCAGATTCTCTGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleValLysPheLeuGlu 200
Qy 666 GTGTCCTGTGCAAGGCCAGCTGGAAGCCCTGACGAGCACTGCCACCACTGGTG--- 722
Db 201 LysAspIleSerGluGluValLeuAsnLysIleIleHisHisThrSerPheAspValMet 220
Qy 723 -----GACCAGTGTGCAAGCTGAGGCCCTGGCC----- 752
Db 221 LysGlnAsnProMetAlaAsnTyrThrLeuProSerSerIleMetAspHisSerIle 240
Qy 753 -----GTGGCCGGGAGAGTGGGTGGAGGACATCTTCACCGTCTCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
Qy 804 AATGAGAAAGTTGACTTGGTGTATAAACAGAAAGATGGAAAGTGTGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrArgLysLysMetAlaGlySerThrIleThrPhe 279

RESULT 30
US-10-199-334-10
; Sequence 10, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-334-10

Alignment Scores:
Pred. No.: 2,95e-31 Length: 283
Score: 408.50 Matches: 88
Percent Similarity: 52.0% Conserv: 57
Best Local Similarity: 31.5% Mismatches: 115
Query Match: 9.3% Indels: 19
DB: 2 Gaps: 3

US-10-768-158-1 (1-2419) x US-10-199-334-10 (1-283)

Qy 81 GAGTTCATGCGTGGCGCTCCGCCCTTCTGCGCGGGAAGATGGAGGATGCCCAAC 140
Db 1 GluValAsnGlyIleLeuMetSerLysMetMetSerGluAsnTrpAspLysIleTrpAsn 20
Qy 141 TTCCCGGTGCGCGCCAGCGAGTGTGATCGTCACTACCCCAAGTCCGGCACCAGCTTG 200
Db 21 PheGlnAlaLysProAspAspLeuIleAlaThrTyrAlaLysAlaGlyThrThrTrp 40
Qy 201 CTGAGGAGGTGCTACTTGGTGAGCCAGCGGCGTGCACCCGATGAGATCGGCTTGATG 260
Db 41 ThrGlnGluIleValAspMetIleGlnAsnAspGlyAspValGlnLysCysGlnArgAla 60
Qy 261 AACATCGACGACGATCCCGGTCTGTGAGTACCCACGACCGC-----GGC 305
Db 61 AsnThrTyrAspArgHisProPheIleGluTrpThrLeuProProLeuAsnSerGly 80
```

```
Qy 306 CTGACATCATCAAGAACTGACCTCTCCCGGCTCATCAAGAGCAACCTACCTACCGC 365
Db 81 LeuAspLeuAlaAsnLysMetProSerProArgThrLeuLysThrHisLeuProValGln 100
Qy 366 TTCTGCGCTCTGACCTCCACAATGGAGCTCAAGGTCTATATATGGTTCGCAACCCC 425
Db 101 MetLeuProSerPheTrpLysGluAsnSerGlnIleIleTyrValAlaArgAsnAla 120
Qy 426 AAGGATCTGGTGGTCTTATTATCAGTTCACCGCTCTCTGCGGACCATGAGCTACCGA 485
Db 121 LysAspCysLeuValSerTyrTyrPheSerArgMetAsnLysMetLeuProAspPro 140
Qy 486 GGCACCTTCAAGAAATCTCCGAGGTTTATGATGATTAAGCTGGGCTACGGCTCTCGG 545
Db 141 GlyThrLeuGlyGluTyrIleGluThrPheLysAlaGlyLysValLeuTrpGlySerTrp 160
Qy 546 TTTGAGCACGTGCGAGGAGTTCTGGAGCACCGCATGCTGCAACGCTGCTTTTCTCAAG 605
Db 161 TyrAspHisValLysGlyTrpTrpAspValLysAspLysHisArgIleLeuTyrLeuPhe 180
Qy 606 TATGAGACATGTCATCGGACCTGGTGCAGTGTGGAGCAGCTGGCCAGATTCCTCGGG 665
Db 181 TyrGluAspMetLysGluAspProLysArgGluIleLysLysIleValLysPheLeuGlu 200
Qy 666 GTGTCCTGTGACAAAGCCAGCTGGAAGCCCTGACGAGGACACTGCCACCACTGGTG--- 722
Db 201 LysAspIleSerGluGluValLeuAsnLysIleIleHisHisThrSerPheAspValMet 220
Qy 723 -----GACCAGTCTGCAACGCTGAGGCCCTGCC----- 752
Db 221 LysGlnAsnProMetAlaAsnTyrThrLeuProSerSerIleMetAspHisSerIle 240
Qy 753 -----GTGGCGCGGGAAGAGTGGGTGTGGAAGGACATCTTCACCGTCTCCATG 803
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260
Qy 804 AATGAGAAAGTTGACTTGGTGTATAAACAGAAAGATGGAAAGTGTGACCTCACGTTT 860
Db 261 SerGluAspPheAspGluAspTyrArgLysLysMetAlaGlySerThrIleThrPhe 279

RESULT 31
US-10-199-329-10
; Sequence 10, Application US/10199329
; Patent No. 6953681
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-10

Alignment Scores:
Pred. No.: 2,95e-31 Length: 283
Score: 408.50 Matches: 88
Percent Similarity: 52.0% Conserv: 57
```

[illegible]

Db 221 LysGlnAsnProMetAlaAsnTyrThrThrLeuProSerSerIleMetAspHisSerIle 240  
Qy 753 -----GTGGCGCGGGAAGAGTGGGCTGTGGAAGGACATCTTCACCGTCTCCCATG 803  
Db 241 SerProPheMetArgLysGlyMetProGlyAspTrpLysAsnTyrPheThrValAlaGln 260  
Qy 804 AATGGAAGTTTGACTCGTGTATAAACAGAAAGATGGGAAAGTGTGACCTCACGTTT 860  
Db 261 SerGluAspPheAspGluAspTyrArgLysLysMetAlaGlySerThrIleThrPhe 279  
RESULT 33  
US-09-609-816-6  
; Sequence 6, Application US/09609816  
; Patent No. 6436684  
; GENERAL INFORMATION:  
; APPLICANT: Woodage, Trevor  
; APPLICANT: Wei, Minh Hui  
; APPLICANT: Kodira, Chinappa  
; APPLICANT: Beasley, Ellen  
; APPLICANT: DiFrancesco, Valentina  
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING  
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN  
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL000669PCT  
; CURRENT APPLICATION NUMBER: US/09/609,816  
; CURRENT FILING DATE: 2000-07-03  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: 60/192,408  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: 60/212,725  
; PRIOR FILING DATE: 2000-06-20  
; PRIOR APPLICATION NUMBER: 09/609,816  
; PRIOR FILING DATE: 2000-07-03  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 304  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-09-609-816-6  
Alignment Scores:  
Pred. No.: 4.33e-31 Length: 304  
Score: 407.00 Matches: 94  
Percent Similarity: 51.7% Conservative: 61  
Best Local Similarity: 31.3% Mismatches: 125  
Query Match: 9.2% Indels: 20  
DB: 2 Gaps: 5  
US-10-768-158-1 (1-2419) x US-09-609-816-6 (1-304)  
Qy 21 ATGGCGGAGACGCGGCGGAGACCCAGACCCCGGGGGAGTTCGAG--AGCAAGTAC 77  
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20  
Qy 78 TTCGAGTTCCATGGCGTGGCGTGGCGCCCTTCTGCCCGCGGGAAGATGGAGGATGCC 137  
Db 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTyrTrpGluLysValCys 40  
Qy 138 AACTTCCCGTGGCGGCCAGGACGCTGTGGATGCTGACCTACCCCAAGTCCGGACACG 197  
Db 41 AsnPheGlnAlaLysProAspAspIleLeuAlaThrThrProLysSerGlyThrThr 60  
Qy 198 TTGCTGACGAGGAGTGTCTACTTGTGTAGCCAGGCGCTGACCCCGATGAGATCGGCTTG 257  
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80  
Qy 258 ATGAACATCGACGACGACCTCCCGTCTCGGAG-----TACCCA-----CAGCGG 302  
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100  
Qy 303 GGCTTGGACATCATCAAGAACTGACCTCTCCCGCTCATCAAGAGCACCTGCCCTAC 362  
Db 101 AspLeuGluPheValLeuMetSerSerProGlnLeuIleLysThrHisLeuProSer 120

Qy 363 CGCTTTCTGCOCTCTGACCTCCACAATGGAGACTCCAAAGGTCACTCATCTATATGGCTCGCAAC 422  
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140  
Qy 423 CCAAGATCTGGTGGTGTCTTATATACAGTTCCACCGCTCTCTCGCGACCATGAGCTAC 482  
Db 141 ProlLysAspCysLeuValSerTyrHisPheHisArgMetAlaSerPheMetProAsp 160  
Qy 483 CGAGGCACCTTTCAAGAATCTCGCGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCC 542  
Db 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlyGlySer 180  
Qy 543 TGGTTTGAGCAGCTGCAGGAGTTCTGGGAGCACCGCATGGACTCGAAGCTGCTTTTCTC 602  
Db 181 TrpPheAspHisValLysGlyTyrTipAlaAlaLysAspThrHisArgIleLeuTyrLeu 200  
Qy 603 AAGTATGAAGACATCGGACCTGGGACCTGGTGCAGTATGGTGGAGCAGCTGCCAGATCTCTG 662  
Db 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220  
Qy 663 GGGGTGCTCTGTGACAAAGCCAGCTGGAAGCCCTGACGGAGCACTGCCACCAGCTGTG 722  
Db 221 GluLysThrLeuSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240  
Qy 723 -----GACCAGTGTCTGCAACGCTGAGGCCCTGCC----- 752  
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260  
Qy 753 -----GTGGCGCGGGAAGAGTGGGCTGTGGAAGGACATCTTCACCGTCTCC 800  
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280  
Qy 801 ATGAATGAGAAGTTTGACTTGGTGTATAAACAGAAAGATGGAAAGTGTGACCTCACGTTT 860  
Db 281 MetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300

RESULT 34

US-10-199-334-6  
; Sequence 6, Application US/10199334  
; Patent No. 6905855  
; GENERAL INFORMATION:  
; APPLICANT: Woodage, Trevor et al.  
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING  
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN  
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL000669Div-3  
; CURRENT APPLICATION NUMBER: US/10/199,334  
; CURRENT FILING DATE: 2002-07-22  
; PRIOR APPLICATION NUMBER: 09/609,816  
; PRIOR FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: 60/192,408  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: 60/212,725  
; PRIOR FILING DATE: 2000-06-20  
; PRIOR APPLICATION NUMBER: 09/609,816  
; PRIOR FILING DATE: 2000-07-03  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 304  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-199-334-6

Alignment Scores:  
Pred. No.: 4.33e-31 Length: 304  
Score: 407.00 Matches: 94  
Percent Similarity: 51.7% Conservative: 61  
Best Local Similarity: 31.3% Mismatches: 125  
Query Match: 9.2% Indels: 20  
DB: 2 Gaps: 5



```
Db 221 GluLysThrLeuSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
Qy 723 -----GACCAGTGTGCAACGCTGAGGCGCTGCGCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
Qy 753 -----GTGGCGGGGAGAGATTGGGCTGTGGAAGGACATCTTCACCGTCTCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
Qy 801 ATGAATGAGAAGTTGACTTGGTGATATAACAGAAAGATGGGAAAGTGAGCACTCACGTTT 860
Db 281 MetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300
RESULT 36
US-10-199-330-6
; Sequence 6, Application US/10199330
; Patent No. 6967094
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-6
Alignment Scores:
Pred. No.: 4,33e-31 Length: 304
Score: 407.00 Matches: 94
Percent Similarity: 51.7% Conservative: 61
Best Local Similarity: 31.3% Mismatches: 125
Query Match: 9.2% Indels: 20
DB: 2 Gaps: 5
US-10-768-158-1 (1-2419) x US-10-199-330-6 (1-304)
Qy 21 ATGGCGGAGACGAGCGCGAGACCCCGGAGGAGTTCGAG--AGCAAGTAC 77
Db 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy 78 TTCGAGTTCCATGGCGTGTGGCTGCGGCTTCTCGCGCGGGAAGATGGAGGAGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys 40
Qy 138 AACTTCCGGTGGCGGCCAGGAGGTGTGGATGTGACCTACCCCAAGTCCGGGACACAGC 197
Db 41 AsnPheGlnAlaLysProAspAspIleAlaThrTyPrLysSerGlyThrThr 60
Qy 198 TTGCTGCAGGAGGTGGTCTACTTGTGTAGCCAGGCGCTGACCCCGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
Qy 258 ATGAACATCGACGAGCTCCCGGTCTCGGAG-----TACCCA-----CAGCCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
Qy 303 GGCCTGGACATCATCAAGAACTGACCTCTCCCGCCTCATCAAGAGCCACTGCCCTAC 362
```

```
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
Qy 363 CGCTTTTTCGCTCTGACCTCCCAACATGGAGATCCCAAGGTCACTATATATGGCTCGCAAC 422
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyValAlaArgAsn 140
Qy 423 CCAAGGATCTGGTGGTGTCTTATATCATAGTTCCACCGCTCTCTCGGACCATGAGTAC 482
Db 141 ProLysAspCysLeuValSerTyrTrpHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACCTTCAAGAAATCTCGCGGAGGTTTATGAATGATGAAGCTGGGCTACGGCTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyGluLysPheMetSerGlyLysValGlyGlySer 180
Qy 543 TGGTTTTCGACGACGTGCGGAGTTCCTGGGAGCACCGCATGGACTCGAACGTCTTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspThrHisArgIleLeuTyrLeu 200
Qy 603 AAGTATGAAGACATGTCATCGGACCTGTGACGATGTGTGGAGCAGCTGGCCAGATTCTGT 662
Db 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy 663 GGGTGTCTGTGACAAAGGCCAGCTGGAAGCCCTGACGGAGCACTGCCACCAGCTGTG 722
Db 221 GluLysThrLeuSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
Qy 723 -----GACCAGTGTCTGCAACGCTGAGCGCTGCGCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
Qy 753 -----GTGGCGCGGGGAAGAGTGGGTGTGGAGAGACATCTTCACCGTCTCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
Qy 801 ATGAATGAGAAGTTGACTTGGTGATATAACAGAAAGATGGGAAAGTGAGCTCACGTTT 860
Db 281 MetAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300
RESULT 37
US-09-609-816-4
; Sequence 4, Application US/09609816
; Patent No. 6436684
; GENERAL INFORMATION:
; APPLICANT: Woodage, Trevor
; APPLICANT: Wei, Minh Hui
; APPLICANT: Kodira, Chinappa
; APPLICANT: Beasley, Ellen
; APPLICANT: DiFrancesco, Valentina
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669PCT
; CURRENT APPLICATION NUMBER: US/09/609,816
; CURRENT FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: HUMAN
US-09-609-816-4
Alignment Scores:
Pred. No.: 1-7e-30 Length: 304
Score: 401.00 Matches: 93
Percent Similarity: 51.7% Conservative: 62
```

```
Best Local Similarity: 31.0% Mismatches: 125
Query Match: 9.1% Indels: 20
DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x US-09-609-816-4 (1-304)

QY 21 ATGCGGAGAGCGAGCGGAGACCCAGCAGCCCGGGGGAGTTCGAG---AGCAAGTAC 77
DB 1 MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20

QY 78 TTCGAGTTCCATGCGGTGCGGCTGCGCCCTTCTGCGCGGGAAGATGGAGGATCGCC 137
DB 21 MetGluValAspGlyValProThrLeuLeuSerLysGluTrpTpgLysValCys 40

QY 138 AACTTCCGGTGGCGGCGGAGCGAGTGTGGATCGTACCTACCCCAAGTCGGGACAGC 197
DB 41 AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTyrProLysSerGlyThr 60

QY 198 TTGCTGCGAGGAGTGTCTACTTGTGTGAGCGGCGCTGACCCGATGAGATCGGCTTG 257
DB 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80

QY 258 ATGAACATCGACGAGCGAGTCCCGTCTCGAG-----TACCCA-----CAGCCG 302
DB 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysGluLysPro 100

QY 303 GGCCTGGACATCATCAAGGAAGTACCTCTCCCGCTCATCAAGAGCCAGCTGCCCTAC 362
DB 101 AspLeuGluPheValLeuGluMetSerProGlnLeuIleLysThrHisLeuProSer 120

QY 363 CGCTTTCTGCGCTCTGACCTCCACAATGGAGACTCCCAAGGTCACTATATGGCTCGCAAC 422
DB 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140

QY 423 CCCAAGGATCTGGGTGCTTATTATTCAGTTCCACCGCTCTCTCGGACCATCAGGTAC 482
DB 141 ProLysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSerPheMetProAsp 160

QY 483 CGAGGCACCTTCAAGATTTCTGCGGAGGTTTATGAATAGTACCTGGGTACGGCTCC 542
DB 161 ProGlnAsnLeuGluPheTyrGluLysPheMetSerGlyLysValValGlySer 180

QY 543 TGGTTTGGACGTCGAGGAGTCTCTGGAGCACCGCATGGACTCGAAGCTGCTTTTCTC 602
DB 181 TrpPheAspHisValLysGlyTrpTrpAlaLysAspMetHisArgIleLeuTyrLeu 200

QY 603 AAGTATGAAGACATGATCGGACCTGTGTGACGATGGTGGAGCAGCTGGCCAGATTCTGT 662
DB 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220

QY 663 GGGGTGCTCTGTGACAGCCCGAGTGGAGCCCTGACGCGGAGCAGTGCACCGCTGGT 722
DB 221 GluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240

QY 723 -----GACCAGTGTGCAACCTGAGGCGCTGCGCC-----752
DB 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260

QY 753 -----GTGGCGGGGAGAGTGTGGCTGTGGAAGGACATCTTCCCGTCTCC 800
DB 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280

QY 801 ATGAATGAGAGTTGACTTGTGTATAAACAGAGAGATGGGAAAGTGTGACCTCACGTTT 860
DB 281 LeuAsnGluAsnPheAspLysHisThrGluLysLysMetAlaGlySerThrLeuAsnPhe 300

RESULT 38
US-10-199-334-4
; Sequence 4, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
```

```
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-199-334-4
```

```
Alignment Scores:
Pred. No.: 1,7e-30 Length: 304
Score: 401.00 Matches: 93
Percent Similarity: 51.7% Conservative: 62
Best Local Similarity: 31.0% Mismatches: 125
Query Match: 9.1% Indels: 20
DB: 2 Gaps: 5
```

US-10-768-158-1 (1-2419) x US-10-199-334-4 (1-304)

```

Db      201  PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy      663  GGGGTGTCTGTGACAAAGCCAGCTGGAAGCCCTGACGGAGCACTGCCAGCAGCTGGTG 722
Db      221  GluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
Qy      723  -----GACCAGTGTGCAACGCTGAGGCCCTGGCC----- 752
Db      241  MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
Qy      753  -----GTGGGCGGGGGAAGAGTGTGGCTGTGGAAGGACATCTTCACCGTCTCC 800
Db      261  IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
Qy      801  ATGAATGAGAAGTTTGACTTGTGTATAAACAGAGATGGGAAAGTGTGACCTCACGTTT 860
Db      281  LeuAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300

RESULT 39
US-10-199-329-4
; Sequence 4, Application US/10199329
; Patent No. 6953681
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-329-4

Alignment Scores:
Pred. No.: 1.7e-30 Length: 304
Score: 401.00 Matches: 93
Percent Similarity: 51.7% Conservative: 62
Best Local Similarity: 31.0% Mismatches: 125
Query Match: 9.1% Indels: 20
DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x US-10-199-329-4 (1-304)
Qy      21  ATGGCGGAGCGGCGGAGACCCCGAGCACCCTCGGGGGAGTTCGAG---AGCAAGTAC 77
Db      1  MetAlaLysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy      78  TTCGAGTTCATGGGTGGCGGCTGCGGCCCTTTCGCGCGGAGAGATGGAGGATCGCC 137
Db      21  MetGluValAspGlyValProThrLeuIleLeuSerLysGluIleTrpGluLysValCys 40
Qy      138  AACTTCCCGGTGGCGCGGAGCGAGCTGTGGATCGTCACCTACCCCAAGTCCGGCAGCAGC 197
Db      41  AsnPheGlnAlaLysProAspAspLeuLeuAlaThrTyrProLysSerGlyThrThr 60
Qy      198  TTGTCGAGGAGGTGGTCTACTTGTGTGAGCCAGGCGCGCTGACCCCGATGAGATCGGCTTG 257
Db      61  TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
Qy      258  ATGAACATCGACGAGCAGCTCCCGTCTCGAG-----TACCCA-----CAGCCG 302

```

```

Db      81  AlaGlnThrLeuAspArgHisAlaPheLeuGluLeuLysPheProHisLysGluLysPro 100
Qy      303  GGCTTGACATCATCAAGGAACCTGACCTCTCCCGCTCATCAAGAGCCACCTCCCTTAC 362
Db      101  AspLeuGluPheValLeuGluMetSerProGlnLeuIleLysThrHisLeuProSer 120
Qy      363  CGCTTTTCGCCCTCTGACCTCCACAATGGAGACTCCCAAGGTCACTATATGGCTCGCAAC 422
Db      121  HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTyrValAlaArgAsn 140
Qy      423  CCCAAGCATCTGGTGGTGTCTTATATATACAGTTCACCGCTCTCTCGCGGACCATGAGTAC 482
Db      141  ProLysAspCysLeuValSerTyrTyrHisPheHisArgMetAlaSerPheMetProAsp 160
Qy      483  CGAGGCACTTTCAGAAATCTGCGGAGGTTTATGAATGATAGCTGGGCTACCGGTCC 542
Db      161  ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlySer 180
Qy      543  TGGTTTGAGCAGCTGCGAGGAGTTCCTGGAGACCCGATGGAAGTCAAGCTGCTTTTCTC 602
Db      181  TrpPheAspHisValLysGlyTrpTrpAlaAlaLysAspMetHisArgIleLeuTyrLeu 200
Qy      603  AAGTATCAACATGTCATCGGACCTGGTGACGATGTGGAGCAGCTGGCCAGATTCTGTG 662
Db      201  PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy      663  GGGGTGTCTGTGACAAAGGCCAGCTGGAAGCCCTGACGGAGCACTGCCACGCTGGTG 722
Db      221  GluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
Qy      723  -----GACCAGTCTGCAACGCTGAGGCCCTGGCC----- 752
Db      241  MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
Qy      753  -----GTGGGCGGGGAGAGTGGGCTGTGGAGGACATCTTCACCGTCTCC 800
Db      261  IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
Qy      801  ATGAATGAGAAGTTTGAATTTGTATATAACAGAAAGATGGAAAGTGTGACCTCACGTTT 860
Db      281  LeuAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300

RESULT 40
US-10-199-330-4
; Sequence 4, Application US/10199330
; Patent No. 6967094
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-1
; CURRENT APPLICATION NUMBER: US/10/199,330
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-330-4

Alignment Scores:
Pred. No.: 1.7e-30 Length: 304

```

```
Score: 401.00 Matches: 93
Percent Similarity: 51.7% Conservative: 62
Best Local Similarity: 31.0% Mismatches: 125
Query Match: 9.1% Indels: 20
DB: 2 Gaps: 5

US-10-768-158-1 (1-2419) x US-10-199-330-4 (1-304)
Qy 21 ATGCGGAGAGCGGCGGAGACCCAGCCAGCACCCCGGGGAGTTCGAG---AGCAAGTAC 77
Db 1 MetAlaIysIleGluLysAsnAlaProThrMetGluLysLysProGluLeuPheAsnIle 20
Qy 78 TTCAGTTCATCGCGTGGCGCTCGCGCCCTTCGCGCGGGAAGATGGAGGATCGCC 137
Db 21 MetGluValAspGlyValProThrLeuIleLeuSerLysGluTrpTrpGluLysValCys 40
Qy 138 AACCTTCGCGTGGCGGCGGAGAGTGTGGATCTGCACCTACCCCAAGTCCGGCACCAGC 197
Db 41 AsnPheGlnAlaLysProAspAspLeuIleLeuAlaThrTyrProLysSerGlyThrThr 60
Qy 198 TTGCTGAGGAGGTGCTACTTGTGAGCAGCGGCGTGCACCCGATGAGATCGGCTTG 257
Db 61 TrpMetHisGluIleLeuAspMetIleLeuAsnAspGlyAspValGluLysCysLysArg 80
Qy 258 ATGACATCGACGACGCTCCGCTCTGGAG-----TACCCA-----CAGCG 302
Db 81 AlaGlnThrLeuAspArgHisAlaPheLeuGluLysPheProHisLysGluLysPro 100
Qy 303 GGCCTGGACATCATCAAGGAACCTACCTCTCCCGCTCATCAAGAGCACCTGCCCTAC 362
Db 101 AspLeuGluPheValLeuGluMetSerSerProGlnLeuIleLysThrHisLeuProSer 120
Qy 363 CGCTTTCGCTCTGACCTCCACAATGGAGACTCCCAAGTCACTATATATGGTCGCAAC 422
Db 121 HisLeuIleProProSerIleTrpLysGluAsnCysLysIleValTrpValAlaArgAsn 140
Qy 423 CCCAGGATCTGGTGTCTTATATCAGTTCCACCGCTCTCTCGGACCATCAGCTAC 482
Db 141 ProLysAspCysLeuValSerTyrTrpHisPheHisArgMetAlaSerPheMetProAsp 160
Qy 483 CGAGGCACTTCAAGAAATTCGCGGAGGTTTATGAATGATAAGCTGGCTACGGCTCC 542
Db 161 ProGlnAsnLeuGluGluPheTyrGluLysPheMetSerGlyLysValValGlySer 180
Qy 543 TGGTTTGAACGTCGAGGAGTTCGGGAGCACCGCATGGACTGCAACGCTCTTTCTC 602
Db 181 TrpPheAspHisValLysGlyTrpTrpAlaLysAspMetHisArgIleLeuTyrIleu 200
Qy 603 AAGTATGAAGACATCATCGGACCTGGTACGATGGTGGAGCAGCTGGCCAGATTCCTG 662
Db 201 PheTyrGluAspIleLysLysAsnProLysHisGluIleHisLysValLeuGluPheLeu 220
Qy 663 GGGGTGCTCTGTGACAGGCCAGCTGGAAGCCCTGACGGAGCACTGCCACAGCTGGTG 722
Db 221 GluLysThrTrpSerGlyAspValIleAsnLysIleValHisHisThrSerPheAspVal 240
Qy 723 -----GACCACTGCTGCAACGCTGAGGCCCTGCGCC----- 752
Db 241 MetLysAspAsnProMetAlaAsnHisThrAlaValProAlaHisIlePheAsnHisSer 260
Qy 753 -----GTGGCGCGGGAAGAGTGTGGCTGTGGAAGGACATCTTCACGCTGCC 800
Db 261 IleSerLysPheMetArgLysGlyMetProGlyAspTrpLysAsnHisPheThrValAla 280
Qy 801 ATGAATGAGAAGTTTGACTTGTGTATAAACAGAGATGGGAAAGTGTCACCTCAGTTT 860
Db 281 LeuAsnGluAsnPheAspLysHisTyrGluLysLysMetAlaGlySerThrLeuAsnPhe 300
```

## RESULT 41

```
US-09-949-016-6526
; Sequence 6526, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6526
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-6526
```

## Alignment Scores:

```
Pred. No.: 1,74e-28 Length: 350
Score: 381.00 Matches: 94
Percent Similarity: 51.8% Conservative: 52
Best Local Similarity: 33.3% Mismatches: 108
Query Match: 8.7% Indels: 28
DB: 2 Gaps: 7
```

US-10-768-158-1 (1-2419) x US-09-949-016-6526 (1-350)

```
Qy 54 CCGGGGAGTTCGAGACCAAGTACTTCGAGTTCATCGGCTCGGCTCGCCCTTCTGC 113
Db 13 ProGlyGlu-----TyrPheArgTyrLysGlyVal-----ProPhePro 25
Qy 114 CGCGGG-----AAGATGGAGGAGATCGCC-----AACTTCCCGGTGCGGCC 155
Db 26 ValGlyLeuTyrSerLeuGluSerIleSerLeuAlaGluAsnThrGlnAspValArgAsp 45
Qy 156 AGCAGCTGTGGATCGTCACCTACCCCAAGTCCGGCACCAGCTGTCGAGGAGGTGTC 215
Db 46 AspAspIlePheIleIleThrTyrProLysSerGlyThrTrpMetIleGluIle 65
Qy 216 TACTTGTGTGACCCAGGCGCTGACCCCGATGAGATCGGCTTGATGAACATCGACGAGCAG 275
Db 66 CysLeuIleLeuLysGluLysAspProSerTrpIleArgSerValProIleTrpGluArg 85
Qy 276 CTCCCGTCTCTGGAGTACCACAGCCGGCTGGACATCATCAAGGAACCTGCTCTCCC 335
Db 86 AlaProTrpCysGlu---ThrIleValGlyAlaPheSerLeuProAspGlnTyrSerPro 104
Qy 336 CGCCTCATCAAGACCCACCTGCTTCTGCGCTTCTGCGCTTCTGACCTCCACATGGAGAC 395
Db 105 ArgLeuMetSerSerHisLeuProIleGlnIlePheThrLysAlaPhePheSerSerLys 124
Qy 396 TCCAAAGTCTATATATGCTCGCAACCCCAAGGATCTGGTGTGTTCTTATTATCAGTTC 455
Db 125 AlaLysValIleTyrMetGlyArgAsnProArgAspValValSerLeuTyrHisTyr 144
Qy 456 CACCGCTCTCTGCGGACCATGAGCTACCGAGGCACTTTTCAAGAAATTCGCGGAGGTTT 515
Db 145 SerLysIleAlaGlyGlnLeuLysAspProGlyThrProAspGlnPheLeuArgAspPhe 164
Qy 516 ATGAATATAAGCTGGCTACGCTGCTGCTTTCGAGCAGCTGAGGAGTTCCTGGGAGCAC 575
Db 165 LeuLysGlyGluValGlnPheGlySerTrpPheAspHisIleLysGlyTrpLeuArgMet 184
Qy 576 CGCATGCACTCGAAGCTGCTTTTCTCAAGTATGAAGACATGCATCGAGCACTTGTGACG 635
Db 185 LysGlyLysAspAsnPheLeuPheIleThrTyrGluGluGlnGlnAspLeuGlnGly 204
Qy 636 ATGTGGAGCAGCTGGCCAGATTCCTGGGGGTGTCTGTGTGACAAGGCCACCTGGAAGCC 695
Db 205 SerValGluArgIleCysGlyPheLeuGlyArgProLeuGlyLysGluAlaLeuGlySer 224
```



```
QY 696 CTGACGGAGCAC-----TGCCACAGCTGGTGGACCAAGTGTCTGCAACGCTGAGGCC 746
   ::::
Db 225 ValValAlaHisSerThrPheSerAlaMetLysAlaAsnThrMetSerAsnTyrThrLeu 244
   ::::
QY 747 CTGCCC-----GTGGCCGGGAGAGATTGGG 773
   ::::
Db 245 LeuProProSerLeuLeuAspHisArgGlyAlaPheLeuArgLysGlyValCysGly 264
   ::::
QY 774 CTGTGGAAGGACATCTCCACGCTCCCATGAATGAGAAATTTGACTTGGTGTATATAACAG 833
   ::::
Db 265 AspTrpLysAsnHisPheThrValAlaGlnSerGluAlaPheAspArgAlaTyrArgLys 284
   ::::
QY 834 AAGATG 839
   ::::
Db 285 GlnMet 286

RESULT 42
US-09-949-016-11211
; Sequence 11211, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11211
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11211

Alignment Scores:
Pred. No.: 1,82e-28 Length: 377
Score: 381.00 Matches: 94
Percent Similarity: 51.8% Conservative: 52
Best Local Similarity: 33.3% Mismatches: 108
Query Match: 8.7% Indels: 28
DB: 2 Gaps: 7

US-10-768-158-1 (1-2419) x US-09-949-016-11211 (1-377)
QY 54 CCGGGGAGTTCGAGACGACGAGTCTTCAGTTCCATGCGGTGCGGCTGCCGCCCTTCTGC 113
   ::::
Db 40 ProGlyGlu-----TyrPheArgTyrLysGlyVal-----ProPhePro 52
   ::::
QY 114 CCGCGG-----AAGATGGAGAGATCGCC-----AACTTCCCGTCCGCGGCC 155
   ::::
Db 53 ValGlyLeuTyrSerLeuGluSerIleSerLeuAlaGluAsnThrGlnAspValArgAsp 72
   ::::
QY 156 AGCAGCTGTGGATCGTCACCTACCCCAAGTCCGGCACCAGCTTGTCTGCAGGAGGTGTC 215
   ::::
Db 73 AspAspIlePheIleIleThrTyrProLysSerGlyThrTrpMetIleGluIle 92
   ::::
QY 216 TACTTGTGTGACCGGGCGTGCACCCCGATGAGATCGGCTTGATGAATCAGCAGCAGCAG 275
   ::::
Db 93 CysLeuIleLeuLysGluGlyAspProSerTrpIleArgSerValProIleTrpGluArg 112
   ::::
QY 276 CTCCCGGTCTGTGGATACCCACAGCCGCGCTGCACATCATCAAGGAAGTCACTCTCC 335
   ::::
Db 113 AlaProTrpCysGlu-----ThrIleValGlyAlaPheSerLeuProAspGlnTyrSerPro 131
   ::::
QY 336 CGCTCTCATCAAGACCCACCTGCCCTACCGCTTTCTGCCCTCTGCACCTCCCAATGGAGAC 395
```

```
Db 132 ArgLeuMetSerSerHisLeuProIleGlnIlePheThrLysAlaPhePheSerLys 151
   ::::
QY 396 TCCAAGGTCACTATATGGCTCGCAACCCCAAGGATCTGGTGGTCTTATTATCAGTTC 455
   ::::
Db 152 AlalysValIleTyrMetGlyArgAsnProArgAspValValSerLeuTyrHisTyr 171
   ::::
QY 456 CACCGCTCTCTCGCGACCATGAGCTACCGAGCCACCTTTCAAGAATCTCGCGGAGTTT 515
   ::::
Db 172 SerLysIleAlaGlyGlnLeuLysAspProGlyThrProAspGlnPheLeuArgAspPhe 191
   ::::
QY 516 ATGAATGATAGCTGGCTACGGCTCTGGTTCCTGGTTCAGCAGCTGCAGGAGTTCTCGGAGCAC 575
   ::::
Db 192 LeuLysGlyGluValGlnPheGlySerTrpPheAspHisIleLysGlyTrpLeuArgMet 211
   ::::
QY 576 CGCATGGACTCGAAACGCTTTTCTCAAGTATGAAGACATGCATCGGGACCTGTGTAGC 635
   ::::
Db 212 LysGlyLysAspAsnPheLeuPheIleThrTyrGluGluGlnGlnAspLeuGlnGly 231
   ::::
QY 636 ATGCTGGAGCAGCTGGCCAGATTCTCTGGGGGTGCTCTGTGACAAAGCCCGCCAGTGAAGCC 695
   ::::
Db 232 SerValGluArgIleCysGlyPheLeuGlyArgProLeuGlyLysGluAlaLeuGlySer 251
   ::::
QY 696 CTGACGGAGCAC-----TGCACACAGCTGGTGGACCAAGTGTGCAACGCTGAGGCC 746
   ::::
Db 252 ValValAlaHisSerThrPheSerAlaMetLysAlaAsnThrMetSerAsnTyrThrLeu 271
   ::::
QY 747 CTGCCC-----GTGGCCCGGGGAGAGTTGGG 773
   ::::
Db 272 LeuProProSerLeuLeuAspHisArgGlyAlaPheLeuArgLysGlyValCysGly 291
   ::::
QY 774 CTGTGGAAGGACATCTTCACCGCTCCATGAATGAGAAATTTGACTTGGTGTATATAACAG 833
   ::::
Db 292 AspTrpLysAsnHisPheThrValAlaGlnSerGluAlaPheAspArgAlaTyrArgLys 311
   ::::
QY 834 AAGATG 839
   ::::
Db 312 GlnMet 313

RESULT 43
US-09-949-016-8464
; Sequence 8464, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8464
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8464

Alignment Scores:
Pred. No.: 1,85e-28 Length: 264
Score: 380.00 Matches: 76
Percent Similarity: 57.7% Conservative: 48
Best Local Similarity: 35.3% Mismatches: 87
Query Match: 8.6% Indels: 4
DB: 2 Gaps: 2

US-10-768-158-1 (1-2419) x US-09-949-016-8464 (1-264)
```

```
Qy 72 AAGTACTTCGAGTTCATGGCGTGGCGTCCGCCCTTCTGCGCGGGAAGATGGAGGAG 131
Db 14 LysLeuLysGluValGluGlyThrLeuLeuGlnProAlaThrValAspAsnTrpSerGln 33
Qy 132 ATCCGCCAACTTCCCGGTGCGGCCAGGACGTCGTGGATCGTCACCTACCCCAAGTCGGCG 191
Db 34 IleglnSerPheGluAlaLysProAspAspLeuLeuLeuCysThrTyProLysAlaGly 53
Qy 192 ACCAGCTTGTGCGAGAGGTGCTTACTTGTGTAGCAGGCGCTGACCCCGATGAGATC 251
Db 54 ThrThrTrpIleGlnGluValAspMetIleGluGlnAsnGlyAspValGluLysCys 73
Qy 252 GGCTTGATGACATCGAGCAGCAGCTCCGCTCTGGAGTAC-----CCACAGCGG 302
Db 74 GlnArgAlaIleIleGlnHisArgHisPhePheIleGluTrpAlaArgProGlnPro 93
Qy 303 ---GGCTGGACATCATCAAGAACTGACCTCTCCCGCCTCATCAAGAGCCACTGCGC 359
Db 94 SerGlyValGluLysAlaLysAlaMetProSerProArgIleLeuLysThrHisLeuSer 113
Qy 360 TACCGCTTTCGCGCTGACCTCCAAATGGAGACTCCAAAGTCATCTATATGGCTGCG 419
Db 114 ThrGlnLeuLeuProProSerPheTrpGluAsnAsnCysLysPheLeuTyValAlaArg 133
Qy 420 AACCCCAAGATCGTGGTGTCTTATATCATGTTCCACCGCTCTCTGCGGACCATGAGC 479
Db 134 AsnAlaLysAspCysMetValSerTyThrHisPheGlnArgMetAsnHisMetLeuPro 153
Qy 480 TACCGAGCAGCTTCAAGAACTTCGCGGAGTTTATGAATGATAAGCTGGGCTACGGC 539
Db 154 AspProGlyThrTrpGluGlyTyPheGluThrPheIleAsnGlyLysValValTrpGly 173
Qy 540 TCCTGGTTGAGCAGCTGCGAGGTTCTGGAGCAGCCGATGAGCTCGAAGCTGCTTTT 599
Db 174 SerTrpPheAspHisValLysGlyTrpTrpGluMetLysAspArgHisGlnIleLeuPhe 193
Qy 600 CTCAAGTATGAAGATCATCGGACCTGCGAGCTGTGAGGATGGTGGAGCAGCTGGCAGATC 659
Db 194 LeuPheTyGluAspIleLysAspProLysHisGluIleArgLysValMetGlnPhe 213
Qy 660 CTGGGGGTGCTCTGTGACAGGCCAGCTGGAAGCCCTGACGGAG 704
Db 214 MetGlyLysLysValAspGluThrValLeuAspLysIleValGln 228
```

RESULT 44

```
US-08-852-481-2
Sequence 2, Application US/08852481
Patent No. 5928931
GENERAL INFORMATION:
APPLICANT: Grun Ph.D., Felix
APPLICANT: Buck Ph.D., Jochen
APPLICANT: Hammerling Ph.D., Ulrich
TITLE OF INVENTION: ISOLATION, PURIFICATION AND CLONING OF
TITLE OF INVENTION: RETINOL DEHYDRATASE
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852.481
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
```

```
APPLICATION NUMBER: US 60/017,178
FILING DATE: 09-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/771
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-852-481-2
Alignment Scores:
Pred. No.: 9,56e-24 Length: 350
Score: 333.00 Matches: 88
Percent Similarity: 46.3% Conservative: 64
Best Local Similarity: 26.8% Mismatches: 106
Query Match: 7.6% Indels: 70
DB: Gaps: 10
US-10-768-158-1 (1-2419) x US-08-852-481-2 (1-350)
Qy 57 GGGAGGTTTCAGAGCAAGTACTTCGAGTTCATGGCGTGGCGTCCGCCCTTCTGCGCG 116
Db 28 GlyAlaPheProThrThrTy-----VallysLeuGlyPro-----Lys 40
Qy 117 GGGAGATG-----GAGGAGATCGCAACTTCCCGGTG 149
Db 41 GlyTyMetValTyArgProTyLeuLysAspAlaAlaAsnIleTyAsnMetProLeu 60
Qy 150 CGGCCAGCAGCGTGGATCGTCACCTACCCAGTCCCGGACACGACGCTGCTGCAGGAG 209
Db 61 ArgProThrAspValPheValAlaSerTyGlnArgSerGlyThrThrMetThrGlnGlu 80
Qy 210 GTGCTCTACTTGTGGTCCAGCGGCGCT----- 236
Db 81 LeuValTrpLeuIleGluAsnAspLeuAsnPheGluAlaAlaLysThrTyMetSerLeu 100
Qy 237 -----GACCCCGATGAGATCGGCTTCATGAAC 263
Db 101 ArgTyIleTyLeuAspGlyPheMetIleTyAspProGluLysGlnGluTyAsn 120
Qy 264 -----ATCACGACGACGCTCCCGGTCTCGAGTAC 293
Db 121 AspIleLeuProAsnProGluAsnLeuAspMetGluArgTyLeuGlyLeuLeuGluTy 140
Qy 294 ---CCACAGCGCGGCTGGACATCATCAAGAACTG-----ACCTCTCCCGGCTCATC 344
Db 141 PheSerArgProGlySerSerLeuLeuAlaAlaValProProThrGluLysArgPheVal 160
Qy 345 AAGAGCCACCTGCGCTACCGCTTCTCCCTCTACCTCCACATGGAGATCCCAAGTGC 404
Db 161 LysThrHisLeuProLeuSerLeuMetProAsnMetLeuAsp---ThrValLysMet 179
Qy 405 ATCTATATGGCTCGCAACCCCAAGATCTGGTGTCTTATTATCATAGTTCACCGCTCT 464
Db 180 ValTyLeuAlaArgAspProArgAspValAlaValSerSerPheHisAlaArgLeu 199
Qy 465 CTGCGGACCATGAGTACCGAGGACCTTTCAAGAAATCTCGCGAGGTTTATGAATGAT 524
Db 200 LeuTyLeuLeuAsnLysGlnSerAsnPheLysAspPheTrpGluMetPheHisArgGly 219
Qy 525 AAGCTGGCTACGGCTCTGTTGACACGTCAGGAGTCTCGGAGCACCGCATGAGC 584
Db 220 LeuTyThrLeuThrProTyPheGluHisValLysAlaIleAlaLysArgHisAsp 239
Qy 585 TCGAACGTCGCTTTTCTCAAGTATGAAGACATGATCGGACCTGGGACCTGGTGGAG 644
```

Db 240 ProtenMetLeuPheLeuPheTyrGluAspTyrLeuLysAspLeuProGlyCysIleAla 259  
Qy 645 CAGTGGCCAGATTCTCGGGGTGCTGTGACAAAGGCCAGCTGGAAGCCCTGACGGAG 704  
Db 260 ArgIleAlaAppPheLeuGlyLysLeuSerGluGluGlnIleGlnArgLeuCysGlu 279  
Qy 705 CACTGCCACAGCTGGTGACACAGCTGTGCAACGCTGAGCGCCCTGCC----- 752  
Db 280 HisLeuAen-----PheGluLysPheLysAsnAenGlyAlaValAsnMetGluAspTyr 297  
Qy 753 -----GTGGCCGGGGAAGAGTGGGCTG 776  
Db 298 ArgGluIleGlyIleLeuAlaAspGlyGluHisPheIleArgLysGlyLysAlaGlyCys 317  
Qy 777 TGAAGGACATCTTCACCGTCTCCATGAATGAGAAGTTTGACTTGTGGTGTATAAACAAGAAG 836  
Db 318 TrpArgAspTyrPheAspGluGluMetThrLysGlnAlaGluLysTrpIleLysAspAsn 337  
Qy 837 ATGGGAAGTGTGACCTCACGTTT 860  
Db 338 LeuLysAspThrAspLeuArgTyr 345  
RESULT 45  
US-09-513-999C-6719  
; Sequence 6719, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCE: 59 US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 6719  
; LENGTH: 60  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 12  
; OTHER INFORMATION: Xaa=Glu or Gly  
US-09-513-999C-6719  
Alignment Scores:  
Pred. No.: 3.18e-22 Length: 60  
Score: 313.00 Matches: 58  
Percent Similarity: 98.3% Conservative: 1  
Best Local Similarity: 96.7% Mismatches: 1  
Query Match: 7.1% Indels: 0  
DB: 2 Gaps: 0  
US-10-768-158-1 (1-2419) x US-09-513-999C-6719 (1-60)  
Qy 1467 ATGAGCAAAATGTCTACTAGGTCTTCAGAAATGAGCGTCTCTGCCAGAGACTTCCAG 1526  
Db 1 MetSerArgAenCysSerLeuGlyLeuGlnSer\*\*ArgProSerAlaArgAspPheGln 20  
Qy 1527 CGGGGGCTCCAAAGGCCCAATGACAGGAGCCCGCGGAGCATGTGCTGAGGGAGTCTG 1586  
Db 21 ArgAlaAlaProLysAlaGlnCysArgGlyAlaArgGlyAlaCysAlaGluGlySerLeu 40  
Qy 1587 CCTGGTGAGGCTGGCAGGTGGGAGTCTAATGCAGTCAGGAGCATTTGTCAGTCAGTGGGTG 1646  
Db 41 ProGlyGluAlaGlyArgTrpGluSerAsnAlaValArgSerIleCysMetGlnTrpVal 60  
RESULT 46

US-09-795-926-2  
; Sequence 2, Application US/09795926  
; Patent No. 6555669  
; GENERAL INFORMATION:  
; APPLICANT: Donoho, Gregory  
; APPLICANT: Hilbun, Erin  
; APPLICANT: Turner, C. Alexander Jr.  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Abuin, Alejandro  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Sands, Arthur T.  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Wilganowski, Nathaniel L.  
; APPLICANT: Hu, Yi  
; APPLICANT: Kieke, James Alvin  
; APPLICANT: Potter, David George  
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME  
; FILE REFERENCE: LEX-0144-USA  
; CURRENT APPLICATION NUMBER: US/09/795,926  
; CURRENT FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 60/185,920  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 60/186,558  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: US 60/191,849  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 303  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-795-926-2  
Alignment Scores:  
Pred. No.: 7.18e-21 Length: 303  
Score: 303.50 Matches: 74  
Percent Similarity: 48.7% Conservative: 61  
Best Local Similarity: 26.7% Mismatches: 99  
Query Match: 6.9% Indels: 43  
DB: 2 Gaps: 9  
US-10-768-158-1 (1-2419) x US-09-795-926-2 (1-303)  
Qy 75 TACTTCGAGTTCCATGCGTGGCTGCGGCGGCGGAAAG---ATGGAG 128  
Db 27 PhePheThrTyrGlnGlyIleProTyrProIleThrMetCysThrSerGluThrPheGln 46  
Qy 129 GAGATCGCCAACTTCCCGGTGCGGCCCGCCAGCGATCGTGGATCGTACCTACCCCAAGTCC 188  
Db 47 AlaLeuAspThrPheGluAlaArgHisAspIleValLeuAlaSerTyrProLysCys 66  
Qy 189 GGCACACAGC-----TTGCTGAGGAGTGTGTCTACTTGGTGAGCAGCGGCGCT 236  
Db 67 GlySerAsnTrpIleLeuHisIleValSerGluIleTyrAlaValSerLys----- 84  
Qy 237 GACCCCGATGAGATCGGCTTGATGAACATCGACGAGCGCTCCCGGCTCTGGAGTACC 296  
Db 85 -----LysLysTyrLysTyrPro 90  
Qy 297 CAG---CCGGCGCTGGACATC-----ATCAAGGAATGACC 329  
Db 91 GluPheProValLeuGluCysGlyAspSerGluLysTyrGlnArgMetLysGlyPhePro 110  
Qy 330 TCTCCCGCTCATCAAGAGCCACCTGCCCTACCGCTTCTGCGCTCTGACCTCCACAAT 389  
Db 111 SerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePheGlu 130  
Qy 390 GGAGACTCCAGGTCACTATATGGCTCGCAACCCCAAGGATCTGGTGTCTTATTAT 449  
Db 131 AsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPheLeu 150

```
Qy 450 CAGTTCACCGCTCTCTCGCGACCATGAGCTACCGAGGCACCTTTCAAGAATTCTGCGG 509
Dy 151 HisPheHisAsnAspValProAspIleProSerTyrGlySerTyrAspGluPheArg 170
Qy 510 AGTTTATGATGATAAGCTGGGTACGGCTCCCTGGTTTGACGAGTTCGG 569
Dy 171 GlnPheMetLysGlyGlnValSerTrpGlyArgTyrPheAspPheAlaIleAsnTrpAsn 190
Qy 570 GAGCAGCGCATGGACTCGAAGCGTCTTCTCAAGTATGAAGACATCGCGGACCTG 629
Dy 191 LysHisLeuAspGlyAspAsnValLysPheIleLeuTyrGluAspLeuLysGluAsnLeu 210
Qy 630 GTGACGATGGTGGAGAGCTGCGCAGATTCCTGGGGGTCTCCTGTGACAAAGCCAGCTG 689
Dy 211 AlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPheLeuThrGlyGluGlnIle 230
Qy 690 GAAGCCCTGACGGAGCACTGCACACGAGCTGTGGACAGCTGCTGCAACGCTGAGGCCCTG 749
Dy 231 GlnThrIleSerVal-----GlnSerThrPheGlnAlaMetArgAlaLysSerGln 247
Qy 750 -----CCGCTGGGC-----CGGGAAGAGTTGGGCTGTGGAAG 782
Dy 248 AspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyGluValGlyAspTrpLys 267
Qy 783 GACATCTTCACCGCTCCCATGAATGAGAAGTTTGACTTGGTGTATAAACAG 833
Dy 268 AsnLeuPheSerGluIleGlnAsnGlnGluMetAspGluLysPheLysGlu 284

RESULT 47
US-10-364-774-2
; Sequence 2, Application US/10364774
; Patent No. 692937
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: POLYNUCLEOTIDES ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/10/364,774
; CURRENT FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 303
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-364-774-2

Alignment Scores:
Pred. No.: 7,18e-21 Length: 303
Score: 303.50 Matches: 74
Percent Similarity: 48.7% Conservative: 61
Best Local Similarity: 26.7% Mismatches: 99
Query Match: 6.9% Indels: 43
DB: 2 Gaps: 9
```

```
US-10-768-158-1 (1-2419) x US-10-364-774-2 (1-303)
Qy 75 TACTTCCAGTTCATCGCGTGGCTCGCG---CCCTTCTCGCGCGGAAG---ATGGAG 128
Dy 27 PhePheThrTyrGlnGlyIleProTyrProIleThrMetCysThrSerGluThrPheGln 46
Qy 129 GAGATCCCAACTTCCCGGTGCGGCCCGCCAGCGATGTGGATCGTCACATCAACCAAGTCC 188
Dy 47 AlaLeuAspThrPheGluAlaArgHisAspAspIleValLeuAlaSerTyrProLysCys 66
Qy 189 GGACACAGC-----TTGCTCAGAGAGTGTCTACTTGTGTGAGCCAGGGGCGCT 236
Dy 67 GlySerAsnTrpIleLeuHisIleValSerGluLeuIleTyrAlaValSerLys----- 84
Qy 237 GACCCCGATGAGATCGCGCTTGATGAACATCGACGAGCAGCTCCCGGTCTGGAGTACCCA 296
Dy 85 -----LysLysTyrLysTyrPro 90
Qy 297 CAG---CCGGCGCTGGGACATC-----ATCAAGGAATCGACC 329
Dy 91 GluPheProValLeuGluCysGlyAspSerGluLysTyrGlnArgMetLysGlyPhePro 110
Qy 330 TCTCCCGCTCATCAAGAGCCACCTGCTCCCTACCGCTTCTGCTGCTGACCTCCACAT 389
Dy 111 SerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePheGlu 130
Qy 390 GGAGACTCCCAAGTCACTATATGGCTCGCAACCCCAAGGATCTGGTGTCTTATAT 449
Dy 131 AsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPheLeu 150
Qy 450 CAGTTCCACCGCTCTCTCGCGGACCATGAGCTACCGAGGCACCTTTCAAGAAATTCGCGG 509
Dy 151 HisPheHisAsnAspValProAspIleProSerTyrGlySerTyrAspGluPheArg 170
Qy 510 AGTTTATGAATGATAAGCTGGCTACGGTCTCTGTTGTTGAGCAGCTGCGAGGATTCGG 569
Dy 171 GlnPheMetLysGlyGlnValSerTrpGlyArgTyrPheAspPheAlaIleAsnTrpAsn 190
Qy 570 GAGCAGCGCATGGACTCGAAGCTGCTTTTCTCAAGTATGAAGACATGCATCGGAGCTG 629
Dy 191 LysHisLeuAspGlyAspAsnValLysPheIleLeuTyrGluAspLeuLysGluAsnLeu 210
Qy 630 GTGACGATGGTGGAGCAGCTGGCCAGATTCCTGGGGGTGCTCTGCAAGGCCAGCTG 689
Dy 211 AlaAlaGlyIleLysGlnIleAlaGluPheLeuGlyPhePheLeuThrGlyGluGlnIle 230
Qy 690 GAAGCCCTGACGGAGCACTGCGCACGAGCTGGTGACCGAGTCTGCAACGCTGAGGCCCTG 749
Dy 231 GlnThrIleSerVal-----GlnSerThrPheGlnAlaMetArgAlaLysSerGln 247
Qy 750 -----CCGCTGGGC-----CGGGAAGAGTTGGGCTGTGGAAG 782
Dy 248 AspThrHisGlyAlaValGlyProPheLeuPheArgLysGlyGluValGlyAspTrpLys 267
Qy 783 GACATCTTCACCGCTCCCATGAATGAGAAGTTTGACTTGGTGTATAAACAG 833
Dy 268 AsnLeuPheSerGluIleGlnAsnGlnGluMetAspGluLysPheLysGlu 284

RESULT 48
US-09-795-926-6
; Sequence 6, Application US/09795926
; Patent No. 6555669
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
```













Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 25394  
LENGTH: 369  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-25394

Alignment Scores:  
Pred. No.: 2,17e-11 Length: 369  
Score: 208.00 Matches: 103  
Percent Similarity: 34.5% Conservative: 19  
Best Local Similarity: 29.1% Mismatches: 96  
Query Match: 4.7% Indels: 136  
DB: 2 Gaps: 22

US-10-768-158-1 (1-2419) x US-09-252-991A-25394 (1-369)

```
QY 840 CCATCTTCTGTTTATACCAAGTCAAACTTCTCATTCATGGAGACGG-----TGAAGA 787
D 840 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 45 ProArgProAlaArgProSerArgSerAlaValAlaGlySerArgAlaAlaSerArg 64
QY 786 TGTCTTCCAGACC-----CAACTTCTCCCGGCCCA 754
D 786 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 65 CysProSerAlaAlaThrProCysHisProThrLeuArgArgArgTTPProGlySer 84
QY 753 CGGCGAGCGCTCAGCGTTGAGCACTGGT-----CCA 721
D 753 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 85 Arg-----ProProArgSerSerThrGlyProGlyArgArgProProArgArgPro 102
QY 720 CCAGCTGTGGCAGTGTCCCGTCA-----GGGCTTCCAGCTGGGCTTGT 676
D 720 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 103 ProArgArgSerAlaAlaAlaSerProArgHisArgThrSerAlaProArgGlyProGly 122
QY 675 CACAGGACACC-----CCAGGAATCTGCCAGCTGTCCA-----640
D 675 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 123 ArgArgThrProAlaSerSerSerGluArgProGlyThrAlaProAlaProProArgArg 142
QY 639 -----CCATCGTCAACCGAGTCCCGATGCATGCTCTTCACTACTTGAGAAAAGCAGCTTCG 586
D 639 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 143 ThrAlaProGlyTTPProGlyProSerAla-----152
QY 585 AGTCCAGCTGGTGTCCAGAACTCTCGACGTGTCTCAACACGAGGAGCGGTAGCCAGCT 526
D 585 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 153 -----AlaGlyAlaGlyArg---ProAlaArgArgProAlaArgPheProProPro---168
QY 525 TATCATTAACCTCCCGGAGAAATTTGAAAGTGCCTCGGTAGTCTCATGTGTCGCCCA 466
D 525 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 169 -----ArgArgCys-----173
QY 465 GAGAGCGTGAAGTGAATAAGACACACCATGCTTGG---GGTTGCGAGCCATAT 409
D 465 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 174 ProAlaGlyProGlyProProArgArgProGlyArgProTTPArgGlyThrAlaArgTyr 193
QY 408 AGATGACCTTGGAGTCTCCATTTGGAGGTGAGAGGCGAGAAAGCGGTAGG-----358
D 408 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 194 Arg-----CysArgGlyCysAlaProAlaProArgArgArgSer 207
QY 357 -----GCAGTGGCTCT 346
D 357 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 208 ProAlaIleProAspAlaArgProAlaCysAlaAlaProArgAlaGlyAlaGlyAlaAla 227
```

```
QY 345 TGATAGCGCGGAGAGGTCTAGTTCTTGATGATGTCAGGCGCGGTGGGTACTCCA 286
D 345 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 228 ArgProAlaGlyGlnGlySer---ProAlaArgCysAlaAlaProSer-----242
QY 285 GGACCGGAGGTGTCTGTCATGTTTCATCAAGCGATCTCATCGGGTTCAGCGCTGGC 226
D 285 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 243 ArgProAlaAlaAlaArgSerAlaThrThrGlyArgArgProSerAlaSerArgAlaGly 262
QY 225 TCACCAAGTAGACCACT-----CCTGCAGCAAGCTGTGCGCG 187
D 225 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 263 ArgProAlaArgCysProAlaGluArgAlaThrAlaProAlaAlaHis-----279
QY 186 ACTTGGGTAGGTGACGATCCACACGTCGCTGG-----154
D 186 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 280 -----ArgAlaTTPProProAlaProAlaProAlaAla 291
QY 153 ---GCCGACCGGAAGTTGGGATCTCTCCATCTTCC-----CGCGGAGAAAGGCG 103
D 153 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 292 ProAlaAlaProGlyAlaAlaArgSerArgProAlaGlyProArgArgGlyArgArg---310
QY 102 GCAGCCGACGCCATGGAATCTCGAAGTACTTCTCGAACTCCCGGGGTGCTGGGG 43
D 102 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 311 -----ArgArgArgCysArgSerAlaProAlaGlySerGlyGly 323
QY 42 TCT---CGGCTCGCTCTCGCCATGCGCGCGCGCTCGCGGT 4
D 42 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 324 AsnProArgLysArgArgLeuGlyCysArgArgArgArg 337
RESULT 57
US-09-252-991A-16906
Sequence 16906, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 16906
LENGTH: 263
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16906
Alignment Scores:
Pred. No.: 1.23e-10 Length: 263
Score: 199.50 Matches: 88
Percent Similarity: 33.1% Conservative: 13
Best Local Similarity: 28.9% Mismatches: 85
Query Match: 4.5% Indels: 119
DB: 2 Gaps: 20
US-10-768-158-1 (1-2419) x US-09-252-991A-16906 (1-263)
QY 816 CAAACTTCTCATTCATGGAGACGG-----TGAAGATGCTCTTCCA 775
D 816 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 31 ArgAlaAlaArgCysTTPArgArgTTPProGlyTTPSerGlySerGluAlaProAlaIle 50
QY 774 GCCCAA---CTCTTCCCGGCCCGCCAGGAGG-----CCTCAGCTTGCAGCACTGGT 724
D 774 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 51 AlaProGlyAlaValProTTPProSerArgGlyArgSerProAlaAspCysArgAlaArg 70
QY 723 CCACAGCTGGT-----GGCAGTGTCTCCGTCA 697
D 723 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 71 ProAlaAlaGlySerArgArgAlaAlaArgProArgArgProArgGlyArgGlyArg 90
QY 696 GGGCTTCCAGCTGGGCTTGTTCACAGGACACCCCGGAGATCTGGCCAGCTGTCTCCA 637
```

[illegible]

RESULT 58  
US-09-252-991A-26169  
; Sequence 26169, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

Alignment Scores:			
Pred. NO.:	1.11e-09	Length:	282
Score:	190.00	Matches:	89
Percent Similarity:	37.1%	Conservative:	23
Best Local Similarity:	29.5%	Mismatches:	106
Query Match:	4.3%	Indels:	84
DB:	2	Gaps:	16
US-10-768-158-1 (1-2419) x US-09-252-991A-26169 (1-282)			
QY	816	CAAACTTCTCATTCATCGAGACGGTGAAGATGTCCTTCCACAGCCCAACTCTTCCCGCG	757
DB	10	ArgThrMetArgGlyArgArgArgArgProCysAlaGlyCysAlaArgAlaArgArgGly	29
QY	756	CCACGGCGAGCGCTCAGCGTTGCAGCAGCTGGTCCACAGCTGGTGGCAGTGCTCCGTC	697
DB	30	ProAlaSerArgProArgLysCysSerAlaGlyArgProAlaProGlyGlySerProSer	49
QY	696	GGGCTTCAGCTGGG-----CCTTGTCCACAGGACA	667
DB	50	AlaSerArgHisGlyArgAlaTrpArgArgSerProAlaProProCysSerAlaArg	69
QY	666	CCC-----CCAGGAATCTGGCCAGCT-----GCTCCACCATCGTCA	631
DB	70	ProGlyProAlaArgGlyProGlyArgArgProGlyArgGlyArgThrProLysSerAla	89
QY	630	CCAGGTCCCGATGATGTCCTTCATCTTGCAGAAAAGCA---CGTTCGAGTCCATCGCGT	574
DB	90	ProValSerProAlaProArgHisArgProArgLysAlaArgAlaSerProArgGln	109
QY	573	GCTCCCGAAGTCTCTGCAGTGGCT-----CAACACGAGGCGGTAGCCGACTTATCAT	520
DB	110	GlyProAlaArgProSerArgAlaAsnArgSerAlaAlaSerArgArgProAlaAlaGly	129
QY	519	TCATAAACCTCCGCGAGAACTTCTTGAAGGTGCCTCGGTAGTCTCATGTCGCCAGAGGC	460
DB	130	AlaGlyThrGlyAlaArgThrCysSerArg-----	139
QY	459	GGTGGAACTGATAATAAGACACCCACAGATCCTTGGGGT-----	421
DB	140	-----ProAlaProGlyGlyArgSerArgThrLeuArgAla	151
QY	420	TGCAGGCAATATAGATGACCTTGGAGTCTCCATTGTGGAGTGCAGAGGCGAAGCGGT	361
DB	152	SerAlaProHisArgSerAlaArgProArgTrpThrProGlyThrArg---LysAlaGly	170
QY	360	AGGCGAGT---GGCTCTTGATAGGCGGGGAGAGGTGAGTTCCTTGATGATGTCGAGC	304
DB	171	ArgSerGlyGlnGlyLeuArgProGlyAlaAspArgCysGly-----CysAlaGly	187
QY	303	CCGGCTGGGGTACTCCAGGACCGGGAGTCTCGTCGATGTTTCATCAAGCCGATCTCAT	244
DB	188	GlySerAlaThrThrProAlaProArgSerVal-----	198
QY	243	CGGGGTGAGCGCCTTGGCTCACCAAGTAGA---CCACCTCTGTCAGCAAGTGTGTCGGG	187
DB	199	-----ProProSerProAlaAlaArgTyrProAlaProValAlaGly-----	211
QY	186	ACTTGGGGTAGGTGACGATCCACAGTCCGCTGGGCGGCACCGGGAAGTGGCGATCTCCT	127
DB	212	-----ArgAspArgArgSerProProAlaProGlySer-----	222
QY	126	CCATCTTCCGCGCAGAGGGCGGCAGCCGACGCCATGGAACTCGAAGTACTTGCTCT	67
DB	223	ProAlaSerArg---ArgLeuSerArgThrAspArgProGlyAsnArgAla-----	238
QY	66	CGAACTCCCGGGGTGTGGGGTCTCGGCTCTCGCTCTCCGCCATCGCGCGCGTGC	7
DB	239	-----ProSerProAlaGlyAsnArgGluIleHisProAlaCysGlnProLeuPhe	255
QY	6	CGTGC	1
DB	256	ArgArg	257





```
QY 243 GATGAG---ATCGGCTTGATGAACATCGACGAGCAGCTCCCGGTCTCTGGAGTAC----- 293
Db 123 ThrHisProLeuLeuPhePheAsnProHisSerCysValGlnAenLeuGluTyrLeuTyr 142
QY 294 -----CCACACCGCGGCTGGACATCATCAAGGAACCTGACCTCTCCCGC 338
Db 143 MetGlyArgGluAenThrMetProAspLeuAspMetLeuAenGlu-----SerProArg 160
QY 339 CTCATCAAGACCCACCTGCGCTTCTCGCTTCTGCGCTCTGACCTCCCAATGGAAGATCC 398
Db 161 LeuPheAlaGlyHisIleProTyr-SerLeuLeuProAlaSerValLeuLysSerGlyThr 180
QY 399 AAGTCTATATATGGCTCGCAACCCNAGGATCTGGTGGTCTTATTATCACTCCAC 458
Db 181 LysIleIleAenIleSerArgAsnArgLysSerThrPheValSerPheThrLysPheGly 200
QY 459 CGCTCTCTGCGGACC----- 473
Db 201 AsnLeuIleAenProAspLysLeuLeuAspLysLeuGluLysSerValAspIlePheAlaSer 220
QY 474 ---ATGAGTACCGAGCACCTTTCAAGAAATCTGCGGAGGTTTATGAATGATAAGCTG 530
Db 221 GlyIleSerPheCysGlyProGluTyrAsnPheGlnAlaGluPheThrAsn----- 237
QY 531 GGCTACGGCTCTGGTTTGACACGTCGACGAGTTCTGGGAGCACCGCATCGACTCGAAC 590
Db 238 -----AlaIleSerThrAsnSerAsn 244
QY 591 GTGCTTTTTCACAGTATGAAGACATGATCGGACCTGTGGTGGTGGAGCAGCTG 650
Db 245 LeuLeuLeuLeuSerTyrGluGluMetLeuLysProValGluAenValLysLysLeu 264
QY 651 GCCAGATTCCTGGGGTGTCC---TGTGACAAAGCCCGCTGGAAGCCCTGACGGAGCAC 707
Db 265 AlaGluPheMetGlyCysGlyPheThrAspAspGluGluLysGlnGlyIleValAspGlu 284
QY 708 TGCCACCAAGCTG-----GTGACACAGTGTGCAACGCTGAGGCCCTGCCCGTGGC--- 758
Db 285 IleValLysLeuCysSerPheAspAsnLeuLysAenGlnValAsnLysAenGlySer 304
QY 759 -----CGGGGAAGAGTTGGGCTGTGG 779
Db 305 SerTyrAenSerLysIleAspAsnLysHisPheArgLysGlyGluValArgAspTyr 324
QY 780 AAGGACATCTTCACGCTCTCCATGAATGAGAGTTTGAATTTGATATAAA 830
Db 325 AlaAenTyrLeuThrSerGluMetIleLysLysLeuGluThrAlaGlyLys 341

RESULT 63
US-09-252-991A-25346
; Sequence 25346, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25346
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25346
Alignment Scores:
Pred. No.: 3.63e-09 Length: 252
```

```
Score: 184.50 Matches: 89
Percent Similarity: 35.4% Conservative: 16
Best Local Similarity: 30.0% Mismatches: 98
Query Match: 4.2% Indels: 94
DB: 2 Gaps: 16

US-10-768-158-1 (1-2419) x US-09-252-991A-25346 (1-252)

QY 822 CCAAGTCAAACTTCTCATTTGAGAGCGGTGAAGATGTCTCTCCACAGCCCACTCTTC 763
Db 21 ProGlyGlnArgSerCysArgTyrArg-----Pro-ProProAlaSerAl 36
QY 762 CCQGGCCACAGGGCAGGGCC-----TCAGCCTTGCAGCACTGGT 724
Db 36 a-ArgProProArgArgProArgTyrValAlaProPheHisSerAlaAlaGlnHisAla 56
QY 723 CCACCACTGGTGGCAGTGTCCGCTCAGGGCTTCCAGCTGGGCTTCTCAGACACCC 664
Db 56 erAla-ProGlnGlnLeuAlaValAlaThrProCysAspProAlaCysArgSerAsp 75
QY 663 CCAGGAATCTGGCAGCTGTCCACCATGCTCACCAGTCCCGATGCCGATGATCTTCTACT 604
Db 76 SerAlaAlaSerAlaProAlaProAlaThrAlaProAlaPro-----CysIle----- 91
QY 603 TGAGAAAAGCACGTTCCAGTCCATGCGGTGCTCCC-----AGAACTCTCTGCAGT 553
Db 92 -----ProAlaGlySerProProGlyAlaArgSerGlyLeuArg 104
QY 552 GCTCAAAACAGAGCCGTAGCCAGCTTATCATTTCAATAAACCTCCGGCAGAAATTTCTGAA 493
Db 105 AlaserArgArgSerArgSerProAla----- 113
QY 492 AGTGCTCTCGTAGCTCATGTCGCGACAGCGGTGGAAGTATGATAATAAGACACCA 433
Db 114 -----ProArgSerArgProThrAlaCysArgProProPro 125
QY 432 GATCCTTTGGGTTGCGAGCCATATAGACTGACCTGGAGTCTCCATTTGTGGAGGTCAAGG 373
Db 126 AlaPro-----ProProArgAlaProThrAspAlaProAlaGlyAlaProArg 141
QY 372 GCAGAAACCGGTAGGCGAGTGTCTTTGATGAGCGGGGAGGAGTCACTTCTTGATGA 313
Db 142 SerArgProAlaArg----- 146
QY 312 TGTCCAGCCCGCTGGTGGTACTCCAGAGCCGGAGCTGCTCGTATGTTTCATCAAGC 253
Db 147 CysAlaGlyArgAla---AlaSerProGlyGlyArgSerAlaArgArgThrArgSerPro 165
QY 252 CGATCTCATCGGGTCAAGCCCTGGC-----TCACCAAGTAGACACCACTCTCTGCA 202
Db 166 Arg---ArgArgGlySerAlaProGlyArgArgCysGlyThrSerAlaProGlnProArg 184
QY 201 GCAAGCTGGTCCGCACTTGGGTTAGTGACATCCACAGTCCGCTGGGCGCGACCGGGA 142
Db 185 ThrGlyTyr-----ArgArgAsnAlaArgGlnSerAlaProAsp 197
QY 141 AGTTGGCGATCT-----CCTCCATCTTCCCGCGGAGAGGGCG 103
Db 198 AlaGlyGlySerHisGlyProGlyProAspArgProGlyThrAlaArgArgAlaArgPro 217
QY 102 GCAGCCGACGCCATGGAACCTCGAAGTACTTGTCTCGAACTCCCGCGGGTGTGGGGG 43
Db 218 ProValProAlaProGlyCysHisArgGlyCysSerArg-----ProGly---TrpArg 234
QY 42 TCT-----CGGCTCTCGCTCTCCGCAATGCCG 16
Db 235 SerAlaGlyArgSerAlaProArgProCysProProCysArg 249

RESULT 64
US-09-949-016-9748
; Sequence 9748, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9748
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9748

Alignment Scores:
Pred. No.: 6.68e-09 Length: 481
Score: 183.50 Matches: 97
Percent Similarity: 36.1% Conservative: 26
Best Local Similarity: 28.4% Mismatches: 101
Query Match: 4.2% Indels: 117
DB: 2 Gaps: 21

US-10-768-158-1 (1-2419) x US-09-949-016-9748 (1-481)

Qy 858 ACGTGAGGTACACTTCCCATCTCTGTTATACACCACTCAAACTTCTCATTCATGG 799
Db 1 ThrSerGlyHisHisArgProAlaAlaAala---ThrProThrGlnGlySer----- 16

Qy 798 AGACGGTGAAGATGTCCTTCCACAGCCCACTTCCCGCCGCCACGGCGAGGCCCTCAG 739
Db 17 -----LeuProAspProSerThrAlaArgHis 25

Qy 738 GGTTCAGACAGTGTCCACAGCTGGTGGCAGTCTCGTCGAGGCTTCCAGCTGGGCCT 679
Db 26 GlnArgAlaArgAspProSerAlaGlyAspGlnSerProVal-----Pro 40

Qy 678 TGTACAGGACACCCCGCAGGATCTGGCCAGCT-----GCTCCACCA 637
Db 41 ValThrArgProProPro-----LeuProAlaSerArgSerThrThrArgAlaGluPro 58

Qy 636 TCGTCACAGGT-----CCGATGTCATCTTCCATCTTCCAGAAAGACGTCGAGT 583
Db 59 SerAlaSerGlySerLeuProSerLeuCysLeuHis-----ArgAlaSer 73

Qy 582 CCATGCGGTGCTCCAGAACTCCTGCACGTCTCAAAACAGGAGCGGTAGCCAGCTTAT 523
Db 74 Pro-----ArgProArgThr-----LeuSerLeuGlnArgAlaProAlaTrp 87

Qy 522 CATTCTAAACCTCGGCGAGAATCTTGAAGGTGCTCGGTAGCTATGTCGCGCAGAG 463
Db 88 AlaAlaGlyLeuSerGlyProAlaArgAspProLeuSerSerProGlnLysGlyArg 107

Qy 462 AGCGGT-----GGAACCT-----GATAATAAGACACCAACCCAGAT 430
Db 108 AlaSerValProGlyThrProGlyProProProProAspSerArgGlyAspProGlu 127

Qy 429 CTTTGGGGTTCGAGCCATATAGATGACCTTGGAGTCTCCATTGTGGAGGTTCAGAGGCA 370
Db 128 ProGlyGlyAlaGlyArgProArg-----HisAspCysArgGluArgAla 142

Qy 369 GAAAGCGGTAGGGCAGGTGGCTCTTTGATGAGCGCGGAGAGGTCAAGTTCCTTGATGAT 310
Db 143 AlaAlaAlaAlaAlaGlyGlyAla----- 150

Qy 309 CCAGCGCGGCTGTGGGTACTCCAGGACCGGAGCTGCTCGATGTTTCATCAAGCCGA 250
Db 151 ---GlyTyrArgGluGlyArgAlaGlyGlyAlaAlaSerAlaProCysSerAla----- 167
```

```
Qy 249 TCTCATCGGGGTACGCCCTTGGCTCACCAGTAGACCACTCTCTGCAGCAAGCTGGTGC 190
Db 168 -----GlySerArgProThrAlaThrSerArgSerSerProTrpProSer----- 182
Qy 189 CGGACTTGGGTAGTCACGATCCACACGTCGCTGGGCCCGCA----- 148
Db 183 ArgThr-----ArgProSerAlaAlaSerArgTrpAlaAlaSerThrSerSerPro 200
Qy 147 -----CCGGAAGTTCGGCATCTCTCCATCT----- 121
Db 201 SerAlaSerProSerThrAlaThrThrProLysSerGlyArgThrAlaSerAlaThrThr 220
Qy 120 -----TCCCGGGCAGAGGCGGCGGCGCAGCGCACGC----- 91
Db 221 SerHisSerThrThrAlaSerSerArgSerArgAlaArgProAlaAlaArgValArgAla 240
Qy 90 -----CATGGAACCTCGAAGTACTTGTCTCGAACTCCCGCGGGTGC----- 49
Db 241 ThrThrGlyArgLeuThrProThrArgArgThrCysSerArgAlaAlaAlaSerCysAla 260
Qy 48 -----TGGGGTCTCGGCTCTCTCCGCCA-----TGCGCGCGCGCT 10
Db 261 AlaAlaSerAlaSerSerAlaArgThrSerProProThrArgLeuThrCysThrThrArg 280
Qy 9 CGC 7
Db 281 Arg 281

RESULT 65
US-09-252-991A-30870
; Sequence 30870, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30870
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30870

Alignment Scores:
Pred. No.: 1.85e-08 Length: 326
Score: 178.00 Matches: 97
Percent Similarity: 32.2% Conservative: 17
Best Local Similarity: 27.4% Mismatches: 112
Query Match: 4.0% Indels: 128
DB: 2 Gaps: 22

US-10-768-158-1 (1-2419) x US-09-252-991A-30870 (1-326)

Qy 822 CCAAGTCAAACTTCTCATTCATG----- 799
Db 5 ProThrArgSerAlaArgArgTrpProAlaArgThrArgLeuArgArgGlyThrProGly 24
Qy 798 -----AGACGTGAAGATGCTTCCACAGCCCACTTCCCC 760
Db 25 ArgCysAlaGlyArgThrAlaArgAlaArgCysProGlyAlaArgArgTrpArgPro 44
Qy 759 GGCCCA-----CGGCGAGGCTCAGCGTTGCAGCACTGGTCCA----- 721
Db 45 AlaProGlyGlyTrpAlaArgHisArgProAlaProArgCysAlaArgProProSer 64
```





Db 250 -----AlaSerArgProSerSerAlaAlaCysProMetAlaArgProThrThrTrp--- 266

Qy 694 CCTGACGGAGCAGCTGCCACCAAGCTGGTGGACCAAGTGTGCAACAGCTGAGGCCCTGCCCG 753

Db 267 -----SerThrAlaThrGly---TrpThrSerThrProAlaAlaSerAlaTrpPro 282

Qy 754 TGGCGCGGGGAAGAGTTGGCTGT-----GGAAGACATCTTCACCG 795

Db 283 TrpAlaArgProProAlaAlaCysGlySerAlaArgThrAlaGlyArgThrGlySerSer 302

Qy 796 TCTCCA 801

Db 303 SerPro 304

RESULT 67

US-09-252-991A-28625

; Sequence 28625, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 28625

; LENGTH: 202

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-28625

Alignment Scores:

Pred. No.: 1.96e-08 Length: 202

Score: 176.50 Matches: 77

Percent Similarity: 30.8% Conservative: 9

Best Local Similarity: 27.6% Mismatches: 86

Query Match: 4.0% Indels: 107

DB: 2 Gaps: 14

US-10-768-158-1 (1-2419) x US-09-252-991A-28625 (1-202)

Qy 825 ACACCAAGTCAAACTTCTCATTCAGAGAGGTGAAGATGCTCTCCACAGCCCAACTC 766

Db 6 SerProThrAlaProProArgArgTrpAlaArg----- 16

Qy 765 TTCCCGCGCCACCGGCGAGGCGCTCAGCGTTGCGAGCACTGGTCCACAGCTGGTGGCAGT 706

Db 17 -----ProArgArgSerArgThrArgCysAlaProProArgAlaGlyArgPro 33

Qy 705 GCTCGTCAGGGCTTCCAGCTGGGCTTGTTCACAGACACCCCGAGGAATCTGGCCA--- 649

Db 34 AlaArgAlaAlaProProAlaGlyArgArgArgSerPro-----CysTrpProArg 51

Qy 648 GCTGCTCCACCATGCTCACCAGTCCCGATCGATGCTTCTACTATTGAGAAAAGCAGCT 589

Db 52 ArgCysArgArgThrProGlyProProArgCys-----AlaGly 65

Qy 588 TCGAGTCCATGGGTGCTCCAGAACTCTCGACGTGCTCAACACGAGGC----- 538

Db 66 SerArgProSerGlyArgProArgArgSerArgArgTyrAlaArgSerTrpProGly 85

Qy 537 -----CGTAGCCAGCTTATCATTTATAAACCTCCCGCAGAAATCTTTGAAAGGTGC 487

Db 86 GlyLeuArgArgHisProAla----- 92

Qy 486 CTCGGTAGCTCATGTCCGCGAGAGCGGTGGAACCTGATATAGACACCCAGATCCT 427

Db 93 -----ArgGlyGlySerGlyCysLeuArgGlyThrAlaArg 104

Qy 426 TGGGTTGCGAGCCATATAGATGACCTTGGAGTCTCCATTGTGGAGTCAAGAGGCAGAA 367

Db 105 TrpGlyGlyArgProValArg-----ArgAlaSerGlyArgArgProPro 119

Qy 366 ---AGCGGTAGCGAGGTGGCTCTTGATGAGCGGGAGAGGTCAAGTTCTTGATGATGT 310

Db 120 GlySerGlyArgArgSerGlySer-----GlyProGlyArgArgSerPro----- 134

Qy 309 CCAGCGCGGCTGGGTACTCCAGAGCCGCGAGTCTCGTCGATGTTTCATCAAGCCGA 250

Db 135 -----CysGlyValProGly----- 139

Qy 249 TCTCATCGGGTACGCCCTGGCTCCACCAAGTAGACCACTCTCTGCAGCAAGTGGTGC 190

Db 140 -----ArgProGlySerPro----- 144

Qy 189 CGGACTTGGGTAGGTGACGATCCACAGCTCGCTGGGCGCGCACCG-----CGA 142

Db 145 -----ArgArgTrpArgArgProValArgArgThrGly 155

Qy 141 AGTTGGCGATCTCCTCCA-----TCTTCCCGCGCAGAAAGCGCGCAGCC 97

Db 156 SerArgArgGluSerProArgArgSerValArgArgProAlaGlyArgArgAlaValPro 175

Qy 96 GCACGCCATGAATCGAAGTACTTGTCTCGAACTCCCGCGGGTGTGGGGGTCT 40

Db 176 AlaTrpArgGlyLeuValLeuHisCysSerCysHisValProHisAlaTrpGlyThr 194

RESULT 68

US-09-252-991A-23981

; Sequence 23981, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 23981

; LENGTH: 525

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-23981

Alignment Scores:

Pred. No.: 3.45e-08 Length: 525

Score: 176.50 Matches: 96

Percent Similarity: 34.2% Conservative: 14

Best Local Similarity: 29.8% Mismatches: 107

Query Match: 4.0% Indels: 105

DB: 2 Gaps: 20

US-10-768-158-1 (1-2419) x US-09-252-991A-23981 (1-525)

Qy 786 TGTCTTTCCACAGCCCAACTCTTCCCGCGCCAGCGGCGCTCAGCGTTGCAGCACT 727

Db 11 CysPro-----GlnProSerProGlyArgArgArgSerProLeuArgGlnArgArg 27

Qy 726 GGTCCACAGCGTGTGGCAGTGTCCGTCA-----GGGCTTCCAGCTGGGCCT 679

Db 28 GlyAlaPro-----GlyGlnArgProAlaPheAlaAspProGlyArgProAlaGlyAsn 45

Qy 678 TGTACAGGACACCCCGAGGAATCTGGCCAGCTCTCCACCATCGTCACCAAGTCCCGAT 619

Db 46 ArgGlyAlaArgGluGlnGly-----ProArgAlaGlnArgProAla 59

```
Qy 618 GCATGCTCTCATCTTGAGAAAGACGCTTCGAGTCATCGGTGCTCCAGCACTCCT 559
Db 60 GlnCysArgPheArgLeuProArgAlaAlaHisHisProGlnProArgProArgPro 79
Qy 558 GCAGCT----- 553
Db 80 ThrGlnGlyArgArgSerLeuArgProGlyHisArgThrArgHisProArgArgGlnArg 99
Qy 552 -----GCTCAACACAGAGCCGTAGCCACCTTATCATTTCAATAAC- 511
Db 100 ProValAlaArgHisArgProArgArgProGlyValProArgArgThrGlyProValArg 119
Qy 510 -----TCCGGCAGATTCTTGAAGGTCCCTCGGTAGCTCATGTCGCGCA 466
Db 120 GlyAspProAlaSerAlaArgArgAlaGlyArgAlaGlyAlaArgArgGlyAla 139
Qy 465 GAGAGCGGTGAATGTGTAATAAGACACACCAGATCCTTGGGTTC- 415
Db 140 ArgSerGlyGlyThrGluGlyLys-----CysArgArgGlyGln 152
Qy 414 CCATATAGATGACCTTGGAGTCTCCATTGTGGAGGTGAGGCGAGGCGAGCGGTAGGCGCA 355
Db 153 Pro-----GlyGlnArgAlaAspGlyValArgArg 162
Qy 354 GGTGGCTCTTGATGAGCGGGGAGAGGTTCAGTTCCTTGCATGATGCCAGGCCCGCTGTG 295
Db 163 GlyProPro-----AlaGlyAspArgArgThrPro-----LeuArgProGlyProAlaAla 179
Qy 294 GGTACTCCAGGACCG--GGAGCTGCTCGTCGATGTTTCATCAAGCCGATCTCATCGGGGT 238
Db 180 AlaLeuProGlyProArgProAlaProArgAlaLeuProLeuSerArgProGlyArgGly 199
Qy 237 CAGCGCTGCTGCTCACCAGTAGACCACTCTCTGAGCAAGCTGGTCCGACCTTGGGGT 178
Db 200 ProGlyProGlyArgArgGlnAlaArgProAlaGly-----Gly 212
Qy 177 AGGTGACGATCCACACCTCGTGGCGCGCACCG-----GGAAGTTGG 136
Db 213 Arg-----ArgArgArgAlaGlnProValAlaGlnArgProAlaGlyHisArg 228
Qy 135 CGATCTCTCCATCTTTC- --CGCGCAGAGGCGCGCAGCGCACGCGCATGGAACCTCGA 79
Db 229 GlnAspProProGlyGlnProProAlaArgProAlaAlaGlyAlaArgArgGly--Arg 247
Qy 78 AGTACTGCTCTCGAACTCCCGCGGGGTGCTGGGGGTCTCGGCTCTGCTCCG----- 25
Db 248 SerProGlyGlyArgGlyAspPro-----PheGlyGlyGlnProArgProProGlnAla 265
Qy 24 -----CCATCGCGCGCGCGT-----CGC 7
Db 266 LeuAlaAlaAlaThrValProProAlaAlaProLeuArgLeuArgAlaGlyProGlyArg 285
Qy 6 CGTGGC 1
Db 286 ArgArg 287

RESULT 69
US-09-795-926-8
; Sequence 8, Application US/09795926
; Patent No. 6555669
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Waite, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James Alvin
; APPLICANT: Potter, David George
```

```
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/191,849
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 148
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-795-926-8

Alignment Scores:
Pred. No.: 1.83e-08 Length: 148
Score: 176.00 Matches: 42
Percent Similarity: 47.7% Conservative: 31
Best Local Similarity: 27.5% Mismatches: 50
Query Match: 4.0% Indels: 30
DB: Gaps: 5

US-10-768-158-1 (1-2419) x US-09-795-926-8 (1-148)
Qy 126 GAGGAGATCGCAACTTCCCGGTGCGGCCAGCCAGCAGCTGGATCGTCACCTACCCCAAG 185
Db 8 GlnAlaLeuAspThrPheGluAlaArgHisAspAspIleValLeuAlaSerTyrProLys 27
Qy 186 TCGGCCACACAGC-----TTGCTGCGAGAGGTGCTACTTGGTGAGCCAGGCG 233
Db 28 CysGlySerAsnTrpIleLeuHisIleValSerGluLeuIleTyrAlaValSerLys--- 46
Qy 234 GCTGACCCCGATGAGATCGCGCTTCATGAACATCGACGAGCAGCTCCCGGCTCTGGAGTAC 293
Db 47 -----LysLysTyrLysTyr 51
Qy 294 CCACAG--CGGGCTGGACATC-----ATCAAGGAAGCTG 326
Db 52 ProGluPheProValLeuLeuGluCysGlyAspSerGluLysTyrGlnArgMetLysGlyPhe 71
Qy 327 ACCTCTCCCGCTCATCAAGAGCCACTGCGCCCTACCGCTTCTGCGCTCTGACCTCCAC 386
Db 72 ProSerProArgIleLeuAlaThrHisLeuHisTyrAspLysLeuProGlySerIlePhe 91
Qy 387 AATGAGACTCCAGGTCATCTATATGCTGCGCAACCCAGGATCTGGTGTGCTCTTAT 446
Db 92 GluAsnLysAlaLysIleLeuValIlePheArgAsnProLysAspThrAlaValSerPhe 111
Qy 447 TATCAGTTCCACCGCTCTCTCGCGGACCATGAGCTACCGAGGACCTTTTCAAGAATTCCTGC 506
Db 112 LeuHisPheHisAsnAspValProAspIleProSerTyrGlySerTyrAspGluPhePhe 131
Qy 507 CGGAGGTTTATGAATGATAAGCTGGGCTACGGCTCTCGCTGG 545
Db 132 ArgGlnPheMetLysGlyGlnGluSer---GlyCysTrp 143

RESULT 70
US-10-364-774-8
; Sequence 8, Application US/10364774
; Patent No. 6929937
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
```



Db 152 AlaProProAlaProArgArgProAlaValProGlySerTrpArgCysArgAlaGly 171  
 Qy 135 -----CGATCTCTCCATCTTCC 118  
 Db 172 ProValArgAlaAlaProAlaProGlyArgGlyTyrArgArgSerProProAlaAla 191  
 Qy 117 -----CGCGGAGAGGCGCGCAGCGCA----- 94  
 Db 192 AlaThrAlaProAlaAlaGlyAlaGlyArgThrSerAlaGlyAlaAlaProAlaGlyArg 211  
 Qy 93 -----CGCCATGGAAGTCCGAAGTACTTGC 70  
 Db 212 ProAlaAlaProProAlaAlaAlaProGlyArgAsnArgProGlySerArgArgThrAla 231  
 Qy 69 TCTCGAAGTCCCGGGGTGCTGGGGTCTCGGCTCGCTCCGCGCATGCGCGCGCGT 10  
 Db 232 GlyArgProArgProGlyArgTrpArgThrArgHisArgGlyArgProAlaAlaAlaGly 251  
 Qy 9 CGCGCTGCGC 1  
 Db 252 ArgAsnArg 254  
 RESULT 72  
 US-09-252-991A-27480  
 ; Sequence 27480, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 27480  
 ; LENGTH: 313  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-27480

Alignment Scores:  
 Pred. No.: 2,85e-08 Length: 313  
 Score: 176.00 Matches: 80  
 Percent Similarity: 31.4% Conservative: 11  
 Best Local Similarity: 27.6% Mismatches: 115  
 Query Match: 4.0% Indels: 84  
 DB: 2 Gaps: 11  
 US-10-768-158-1 (1-2419) x US-09-252-991A-27480 (1-313)  
 Qy 753 GGGCAGGGCTCAGCGTTGCAGCACTGGTCCACAGCTGCT---GGCAGTGTCCGTC 697  
 Db 32 ArgAlaGlyProArgArgSerProProGlySerProAlaSerArgGlyCysArgProArg 51  
 Qy 696 GGGCTTCAGCTGGGCTTGTACAGGACACCCCGAGGAATCTGGCCAGCTGTCTCCACCA 637  
 Db 52 ArgGlySerAlaGlyProAsnArgAsnProAlaProGlyArgArgProGlyTyProAla 71  
 Qy 636 ---TCGTACACAGGTCGGATGCATGCTTTCATCTTGAGAAAGACGCTTCAGTCCA 580  
 Db 72 ArgSerAlaProGlyProProAla-----SerAlaAlaAlaSerAlaPro 86  
 Qy 579 TGCGGT---GCTCCAGAACTCCTGCAGT-----GCTCAACACAGG 541  
 Db 87 ProGlyCysValProProAlaAlaSerArgAlaGlyAspTrpProAlaAlaArgAsnAsp 106  
 Qy 540 AGCGGTAGCCAGCTTATCATTCATAACCTCCGCGCAGAAATCTTTGAAGGTGCTCGGT 481  
 Db 107 SerArgArgPro-----HisArgGlyValProGlyArg-----ProGlyCysArgGly 122

Qy 480 AGCTCATGTTCCGACAGAGCGGTGGAACTGATATAAGACACCACAGATCCTTGGGT 421  
 Db 123 AlaProAlaSerAlaSerMetAspSerValSerAla----- 134  
 Qy 420 TGCAGCCATATAGATGACCTTGGAGTCTCCATTTGTGGAGTCCAGAG----- 373  
 Db 135 -----SerGlnHisAlaGlyGlnAlaArgProThrHisArg 146  
 Qy 372 -----GCAGAAAGCGGTAGGCGAGTGGCTTCTTGATGAGCGGGAG 331  
 Db 147 GluAlaProArgAlaAlaProHisSerGlyArgAla----- 158  
 Qy 330 AGTCAGTTCCTTGATGATGTCCAGGCCCGGCTGTGGTACTCCAGGACCGGAGCTGCT 271  
 Db 159 -----AlaAsnProGlyArgGlyAlaAla 166  
 Qy 270 CGTCGATGTTTCATCAAGCCGATCTCATCGGGTCCAGCGCTGGCTCACCAGTAGACCA 211  
 Db 167 ArgProAlaAlaArgProGlySerAlaProAlaAspProProAlaGluSerProAlaPro 186  
 Qy 210 CTTCTCAGCAAGCTGGTCCCGACTTGGGTAGGTGAGTCCACACGTGCTGGGCC 151  
 Db 187 AlaProAlaSerGlyAlaAlaArgProAlaProArgProGlyProArgHisAlaAla 206  
 Qy 150 GCACCGGGAAGT----- 139  
 Db 207 ValAlaGlyAlaArgAspCysProAlaArgGluSerAlaValAlaAlaThrGlyValPro 226  
 Qy 138 -----TGCGCATCTCTCCATCTTCCCGGCGAGAGGGCGGACGCGCACGC 91  
 Db 227 AlaArgArgCysTrpArgArgProAlaProAlaAlaGlyThrGlyAlaAlaProAlaPro 246  
 Qy 90 CATGGAACCTCGAAGTACTTCTCTCGAACTCCCGGGGTGCTGGGGTCTCGGCTCGC 31  
 Db 247 GlyCysProValProAlaGlyGlyArgProProCysTrpArgGlyArgProArg 266  
 Qy 30 TCTCCGCGCATCGCGCGCGCTGCGCGCTCGC 1  
 Db 267 ProGlyAlaHisArgArgProAlaArgLys 276

## RESULT 73

US-09-252-991A-28620  
 ; Sequence 28620, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 28620  
 ; LENGTH: 267  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-28620  
 Alignment Scores:  
 Pred. No.: 3,26e-08 Length: 267  
 Score: 175.00 Matches: 86  
 Percent Similarity: 32.8% Conservative: 13  
 Best Local Similarity: 28.5% Mismatches: 111  
 Query Match: 4.0% Indels: 92  
 DB: 2 Gaps: 17  
 US-10-768-158-1 (1-2419) x US-09-252-991A-28620 (1-267)

```
Qy 783 CTTTCCACAGCCAACTCTTCCCGGCCACCGGCGGCGCTCAGCGTTGCGACACTGGT 724
|||
Db 15 ProSerThrAlaGly-----ProGlyHisProGlyHisThrAlaGlyCysAlaGlyHis 32
|||
Qy 723 CCACAGCTGGTGGCAGTCCCTCCGTCAGGGCTTCCAGCTGGCCCTTGTACAGG---ACA 667
|||
Db 33 ArgGlyArgAlaGlyProAlaProGlyTrpSerPro-----HisArgAlaAla 48
|||
Qy 666 CCCCAGGAATCTGGCCAGCTGCTCCACCATGCTCACCAGGTCCCGATGCATGT----- 613
|||
Db 49 SerProAlaGlyTrpProGlySerThrAlaArgSerProGlyProGlyAlaGlyCys 68
|||
Qy 612 -----CTTCATCTTGAGAAAGACACGTTCCAGTCCATGCGCGTGT 571
|||
Db 69 TrpProGlyArgThrArgHis-----ArgAlaAlaArgGlyThrProAlaAsnTrp 86
|||
Qy 570 CCCAGAACTCTGACAGTCTCAACACGAGGCGGTAGCCAGCTTATCATTTCAAAACC 511
|||
Db 87 ArgArgProProAlaProGlyArg----- 94
|||
Qy 510 TCCGCGCAGAATCTTTAAAGGTGCTCGCTAGCTCATGCTCCGACAGAGCGGTGGAAC 451
|||
Db 95 SerSerArgValArgArgSerCys---SerProProThrSerAlaCysArgProAlaThr 113
|||
Qy 450 GATATAAGACACACACAGATCCTTGG-----GGTTGCGAG--- 415
|||
Db 114 GlyArgHisArgProGlyArgProTrpArgArgArgAlaProProGlyCysAlaArg 133
|||
Qy 414 -----CCATATAGATGACCT-----TGGAGTCTCCATTTGGAGTTCAG 376
|||
Db 134 AlaAlaProAlaGlyProSerArgTrpProArgArgGlySerArgProProGlyGlu 153
|||
Qy 375 AGGCGAGAAAGCGGTAGGCGAGGTGCTTTGATGAGCGCGGAGGTTCAGTTCCTTGA 316
|||
Db 154 ProAlaArgArgGlyCysAlaGlyAlaSer----- 163
|||
Qy 315 TGATGTCAGCCCGCGGTGTTGGGTACTCCAGAACCGGAGTGTCTGCGATGTTCAATCA 256
|||
Db 164 -----ArgArgProArgArgThrAlaProGlySerGly----- 174
|||
Qy 255 AGCGATCTCATCGGGTCAGCGCCTGGC-----TCACCAAGT 217
|||
Db 175 -----ArgProAlaArgProGlyArgThrAlaCysAlaArgAlaAlaAlaThr 190
|||
Qy 216 AGACCACTCTCTGCA-----GCAAGTGGTGGCGGACCTTGGGGT 178
|||
Db 191 ArgProAlaProAlaAlaProAlaGlyArgThrLeuAlaAlaCysArgProArgArg 210
|||
Qy 177 AGGTGACATCCACA-----CGTCGTCGGCCGCGCAGCGGAAAGTTGGCGATCT 130
|||
Db 211 ArgArgArgThrThrThrGlyArgArgArgArgTrpSerAlaProAla-----ArgArg 228
|||
Qy 129 CTTCAATCTCCCGGCGAGAGGGCGCGCAGCCGATGGAATCGAAGTACTTGC 70
|||
Db 229 ProProGlyTrpArgSerGlyThrGlyArgProArgArgThrAlaProProAlaAlaCys 248
|||
Qy 69 TCTCGAACTCCCGCGGTGTGGGGGTCTCGGCTCGCTCTCCGCCATGCGCGCGGT 10
|||
Db 249 ArgArgAlaProArgGlyAla-----ArgArgAlaThrGlyCysArgTrpArg 264
|||
Qy 9 CGCGGT 4
|||
Db 265 ArgArg 266
|||
```

RESULT 74  
US-09-252-991A-24634  
; Sequence 24634, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136

```
; CURRENT APPLICATION NUMBER: US/09/252.991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 24634  
; LENGTH: 320  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-24634  
  
Alignment Scores:  
Pred. No.: 4.55e-08 Length: 320  
Score: 174.00 Matches: 94  
Percent Similarity: 31.0% Conservative: 23  
Best Local Similarity: 24.9% Mismatches: 125  
Query Match: 4.0% Indels: 135  
DB: 2 Gaps: 19  
  
US-10-768-158-1 (1-2419) x US-09-252-991A-24634 (1-320)  
  
Qy 1 CCGAGCGGCGCGCGCGCGCATGCGGAGA-----GCGAGCGCGAGACCC 45  
|||  
Db 5 SerThrAlaProSerThrProCysArgArgProSerAlaSerAlaAlaGlyProArgThr 24  
|||  
Qy 46 CCAGCACCCCGGGGAGTTCGAGAGCAAGTACTTCG-----AGTTCCATGGCGTGGCG 99  
|||  
Db 25 Pro---ProArgThrSerPheSerAlaSerArgProArgValThrSerProThrCysGly 43  
|||  
Qy 100 TGCGCGCCTTCTGCGCGGGAAGATGGAGAGATGCCCAACTTCCCGGTGGCGCCAGCG 159  
|||  
Db 44 AspCysAlaArgSerThrSerArgTrpArgThrSer----- 55  
|||  
Qy 160 ACGTGTGATCGTCACTTACCCCAAGTCCCGCACCCAGCTGCTGTCGAGAGGTGTTACT 219  
|||  
Db 56 -----SerProAlaProSerProCys----- 63  
|||  
Qy 220 TGGTGGAGCGCGCTCACCCTCGATGAGATCGGCTTGATGACATCCAGCAGCAGCTCC 279  
|||  
Db 64 -----AlaSerAlaSerAla----- 68  
|||  
Qy 280 CGGTCTCTGGAGTACCCACAGCGCGGCTGGACATCATCAAGGAACCTGACCTCTCCCGCC 339  
|||  
Db 69 ArgSerGlySerSerArgArgAla-----Cys 78  
|||  
Qy 340 TCATCAAGAGCCACTGCTCCCTTACCGCTTCTGCGCTCTGACCTCCACAAATGGAGACTCA 399  
|||  
Db 79 AlaSerThrAlaSerCysSerProIleSerProAlaCysArgAlaThrAlaProPro 98  
|||  
Qy 400 AGGTCACTATATGGTCCGCAACCCCAAGGATCGTGGTGTCTTATTTATCATGTTCCACC 459  
|||  
Db 99 ArgArgSerProGlyThrAlaProArg----- 108  
|||  
Qy 460 GCTCTCTCGGACCATGAGCTTACCGAGGCACCTTTCAAGAATTTCTCGCGAGGTTTATGA 519  
|||  
Db 109 ---AlaCysThrAlaArgAlaThrSerThrAlaIleArgArgSerAlaCysGlyMet--- 126  
|||  
Qy 520 ATGATAAGCTGGGTACGGCTCTGTTTGTAGACGTGACGAGTTCCTGGGAGCACCGCA 579  
|||  
Db 127 -----ProTrpLysSerAlaMetArgAlaSerSerThrArgArgSerSerSer--- 142  
|||  
Qy 580 TGGACTCGAAGCTGCTTTTCTCAAGTATGAACATGCATCGGAGC-----TGG 630  
|||  
Db 143 -----CysCysAlaAsnThrGlyLeuProTrp 151  
|||  
Qy 631 TGACGATGGTGGAGCAGCTGGCCAGATTCTCTGGGGGTCTCTGTGCAAGGCCAGCTGG 690  
|||  
Db 152 -----TrpTrpProIleProProAlaSerGlyArgCysSerAlaThrSer----- 166  
|||  
Qy 691 AAGCCCTGACGAGCAGCTGCCACCAGCTGGTGGAC-----AGTGTGCAACGCTG 741  
|||  
|||
```

Db 167 AlaProThrSerSerThrCysAlaCysMetAlaThrArgSerSerThrAlaAlaAla 186  
Qy 742 AGGCCCTGCCGTCGGCGGGGAAGAGTTGGCTGTGAAGACATCTTCAACGCTCCA 801  
Db 187 ArgIysArgProTrpAlaSerGlyArgHisGly-----SerAlaProGlyPro 202  
Qy 802 TGAATGAGAAGTTTGACTTGGTGATATAAC-----AGAAGATGGAAAGT 846  
Db 203 AlaAlaThrArgProThrThrCysAlaGlyProAlaArgThrProAlaArg- 219  
Qy 847 GTGACCTCACGTTTGACTTTTATTATAATAACAGAAACAACACCTGCATGCTCA 906  
Db 220 -----ArgArgArgArg-AlaThrCysThrAlaThr-- 229  
Qy 907 ACCCAGACGCTACTAGCCAAAAGTCCTGTATGATCATTTATTCCT 955  
Db 230 -----SerThrThrThrSerArgCysAlaHisProThrProGlyAlaSerCy 246  
Qy 956 ---TGCTGGACAACTCTGGAACGCGTGTGAACACAGCGGGGAAGAGCGCG 1011  
Db 246 sSerCysTrpAenSerThrAlaAlaCysAla----- 256  
Qy 1012 TGAGCGAGGAGTGTGATGATTCCTCAACCGAAGCAGCTGTCTGCGCTT 1060  
Db 257 -----ArgCysProAspSerCysProMetThrCysGlyArgVal 269  
RESULT 75  
US-09-252-991A-28001  
; Sequence 28001, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 28001  
; LENGTH: 657  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-28001

Alignment Scores:  
Pred. No.: 7,8e-08 Length: 657  
Score: 173.50 Matches: 184  
Percent Similarity: 27.1% Conservative: 44  
Best Local Similarity: 21.9% Mismatches: 204  
Query Match: 3.9% Indels: 410  
DB: 2 Gaps: 42

US-10-768-158-1 (1-2419) x US-09-252-991A-28001 (1-657)

Qy 2256 AGAAAGCAGCTCCCTCAAGAAACACAGGCAAGGTGCTGCGAGCCCATCGATGC 2197  
Db 121 ArgArgHisProProCysArgGluArgGlnAla----- 132  
Qy 2196 CACGTGCCACCTGCTGCGCGCAGCTTCTCCCATTAACAAACAA- 2149  
Db 133 ---ValProGlyGlnProAlaAlaAlaGlnLeuGlnAlaArgAlaHisArgLeuPro 151  
Qy 2148 -----CAAGCTTCTCTACACCCGCTTTCAATACAGCACCAAGAACGA 2104  
Db 152 ArgProArgLeuAspArgAlaArgLeuArgGlnArg----- 163  
Qy 2103 GACCCCTTCCACAGCAGCACCCCTCTGACCCACCACTTTGGGTGCTTCCACGCCACATAA 2044  
Db 164 -----SerAlaPro----- 166

Qy 2043 ATGACACACACCCCGTCCCGACGGGTTGGCTGCTACTGGCCCCCGGACTCTCCATCG 1984  
Db 167 ---AlaGlnArgProAspProAlaGly----- 175  
Qy 1983 CTTCAGCGAATGTGGTGCAGAAATATTAGGAGTGAAGCAATATAAGACCTCACCTTTACTC 1924  
Db 176 ArgGlnArgAlaGlyValArgSerLeuCysAlaSerGlyLeuArgThr----- 191  
Qy 1923 CTTCAGATACACGAGGAGTGAAGTGACAGACACGAGCGGCGCGTCCCAACCTGATCC 1864  
Db 192 -----GlyAlaArgHisLeuAsp-Arg-----Pr 199  
Qy 1863 GAGTGAACACGGGTACACTCGCAAAATGGTCTCCACCGCTGAGCGCTCAGGTGCAC 1804  
Db 199 oglyGlnArgProArgArgSerHisProGlyPhePro----- 211  
Qy 1803 AGAACAACTGAATGATAAAGACCCGGGATTTTAGAAATTTTACTATTAAAGCATCTGCTT 1744  
Db 212 -----ArgArgArgAlaHisArgArgLeuLeuSerAla----- 222  
Qy 1743 GCCAGCATTAGTAGGCTATCACCTTCAGAAACAAGTAAGGCTGTGGAATTCAGCTCAA 1684  
Db 223 -GlnArgLeuValGly----- 227  
Qy 1683 TTCCGAGCGCAACTCGGTCCTTTGGTGGCGGACTCTCCACCCACTG----- 1638  
Db 228 ---AlaArgHisProGlyLeuGlyIleProAlaAlaArgProValProValGluGluLe 246  
Qy 1637 -----CATGCAATGCTCTCTGACTGCTGACTGATAGTATCC 1606  
Db 246 uHisHisGlnArgPheAlaLeuGlyHisArgArgSerProGlyAlaValProProAr 266  
Qy 1605 ACCTGCAGCCTCACAGGAGACATTCCTCAGCACATGCT----- 1565  
Db 266 gProAlaSer-AlaProArgArgProAlaAlaAlaAlaValSerLeuArgArgGlnA 286  
Qy 1564 -----CGCGGGCTCTGCTGCTGGGCTTTGGAGCGCGCTGGAAGTCTCTGG 1513  
Db 286 rgProGlyProArgArgLeuArgHisArgAlaGlyGlyProAlaAlaHisArgGlyProA 306  
Qy 1512 CAGAAGGACGTCCTCTGTAAGACCTAGTGAAACAATTTCTGCTCATTTTCAGGAAGTTCT 1453  
Db 306 laArgGluArgPro----- 310  
Qy 1452 GAAGAAGTGAACCTCGTCAGATAAACAAGTGTCAAGATTCCTCCCAACACACCACTC 1393  
Db 311 -----AlaAlaAlaProAspProGluGlnProGluH 323  
Qy 1392 ATCCTCTGTTCTCACCCTATCTCAGAACCAAGTGAGAACAGATGGCTGCTGCTTAGAT 1333  
Db 323 isValLeuAspGlyGlyAlaAspGlyArgProProGlnArgGlnTrpLeuProVal---- 341  
Qy 1332 AATAAACTACAAAAATAAGAGTTAACCTTGAGATCTTTCTACCATTCGGGTGTGGCTC 1273  
Db 342 -----AlaAlaAla 344  
Qy 1272 GCTCTGATTCCTTGGAAATGAACCTTTATTTGTTTACTGACATTTATGTAGATTTC 1213  
Db 344 rgArg-----ProValArgLeu-----ArgHisP 352  
Qy 1212 CAGTGAAGGCTCTATAAATAACAATAATACGGGTTGAAAGGCAGACATTTCTAGTTGC 1153  
Db 352 roValArgArg----- 355  
Qy 1152 ATATATTACAGGCTTTATCTTACGGTCCAGGCCATTGGAACCTGCAATGTGGAGACTGTT 1093  
Db 356 -----ArgAlaArgProValArgLeuAlaAlaGlyAsp---- 366  
Qy 1092 TGTAAATCAGATGGAGA-----GGTGCAGGTTCTAAAGCGGAGACAGCTGCTTC 1042  
Db 367 ---His-ArgTrpArgGlnAlaAlaGlyAlaThrAlaSerArgArgAspProAlaPhe 384

```
Qy 1041 GGTGGGAATCATCATCCCTCCGCTCAGCCGCTCTTCCCTTCCCCCGCTG----- 988
Db 385 -----LeuGlyArgArgArgGlyHisProProAlaProAla 397
Qy 988 ----- 988
Db 398 AlaSerAlaArgAspMetPheArgSerValSerAlaAsnValAlaAlaLysCysSerArg 417
Qy 987 -----TTTCACACGCTGCTTCCA 970
Db 418 ProIleArgGlyArgAlaAspGluAlaLeuTyrLeuLeuProPheHisLeuLeuPro 437
Qy 969 GAGTTTGCAGCAAGGAATAAATGAATGCATACAGACTTTTGGCTAGTACTGTCTG 910
Db 438 GlyAlaHisArgProGlyAlaGlu----- 445
Qy 909 GGTATTGTGAGCATGCAGGTTGTTGTTCTGTTATTATAATAAAGTCAAACTGAGGT 850
Db 446 -----Gly 446
Qy 849 CACACTTTCCCATCTTCTGTTTATACACCAAGTCAAACCTTCTATTCTATGAGACGTGA 790
Db 447 AlaGlyLeuProValAlaAlaGlyGlnProGlyPro-----ArgArgArg 461
Qy 789 AGATGCTCTCCACAGCCCAACTCTTCCCGGCCCAACGG-----GCAGGCCCTCAGCGTTGC 733
Db 462 ArgThrProProThrArg-----LeuProGlyAlaGlnProAlaGlyProArg----- 477
Qy 732 AGCACTGTCCACAGCTGTGGCAGTGTCTCGTCAGGCTTCCAGCTGGGCTTGTCTAC 673
Db 478 ---ThrGly-----ProAlaGlyArgArgGlyArgAlaAlaAspProValAlaGlyAspHis 495
Qy 672 AGGACACCCCGAGGAATCGGCCAGCTCTCCACCA-----TCGTACCA 628
Db 496 ArgValProArgGlyAlaLeuProAlaProArgProAlaLeuGlnArgSerAlaAlaPro 515
Qy 627 GGTCCGATGATGTCTTCATCTACTTGAGAAAGACGTTTCGAGTCCATGCGGTCTCCC 568
Db 516 ArgPro-----ArgAlaArgSerArgGlyAlaGlyArgLeu 527
Qy 567 AGAAGTCTGACGCTCAACACGAGGCGTAGCCAGCTTATCATTCATAAACCTCC 508
Db 528 ArgHisProProValAlaGlnArgGlnArgAlaGlnProAla----- 541
Qy 507 GGCAGAAATCTTGAAAGTGCTCGGTAGCTCTGTCGTCAGGAGCGGTGGAAGTGTAT 448
Db 541 ----- 541
Qy 447 AATAAGACACCCAGATCTCTGGGTTGCGAGCCATATAGATGACCTTGGAGTCTCCAT 388
Db 542 -----AlaProValGlyThrArg----- 547
Qy 387 TGTGGAGTCCAGAGGCGAAGCGGTAGGCGAGGTGCTCTTGATGAGCGGGAGAGG 328
Db 548 -----ArgArgAla----- 550
Qy 327 TCAGTCTCTGATGATGTCAGGCGCGCTGTGGGTACTCCA-----GGACCGGAGCT 274
Db 551 -----GlyProAlaValAspArgProLeuGlyArgProGlySer 563
Qy 273 GCTGTCGATGTTTCATCAAGCCGATCTCATCGGGT-----CAGCGCCCTGGCTCA 223
Db 564 ArgArgArgAlaThrAspArg---ArgProGlyLeuValLeuLeuArgProAlaArg 582
Qy 222 CCAAGTAGA-----CCACCTCTGCGACAGCTGGTCCCGACTTGGGTAGGTGACGA 169
Db 583 ProGlyArgArgLeuProGlyProAlaAlaLeu-----ArgArgArg 596
Qy 168 TCCACACGCTCGCTGGCGCCGACCGGAAGTTGGCGA-----TCCTCTCCATCTTCC 118
Db 597 AlaLeuArgArg-----GlyAlaGlyArgLeuAlaThrAspProProGlyArg 612
Qy 117 CGC---GGCAGAGGGCGGCGCCACGCCCATGGAACCTCGAAGTACTTGTCTCGAAGT 61
```

```
Db 613 ArgProGlyArgArgPro-----SerGlyVal 621
Qy 60 CCCCCGGGTCTGGGGTCTCGGCTCGCTCTCGCCATCGCGCGCGCTCGCGTGC 1
Db 622 ProProGlyProSerGly---GlnProAlaGlyHisProGlyArgLeuArgAspArgArg 640

RESULT 76
US-09-252-991A-31760
; Sequence 31760, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31760
; LENGTH: 1706
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31760

Alignment Scores:
Pred. No.: 1,72e-07 Length: 1706
Score: 172.50 Matches: 106
Percent Similarity: 35.3% Conservative: 26
Best Local Similarity: 28.3% Mismatches: 98
Query Match: 3.9% Indels: 144
DB: 2 Gaps: 24

US-10-768-158-1 (1-2419) x US-09-252-991A-31760 (1-1706)
Qy 762 CCGCGGCCCA-----CGGCGAGGCTCAGCGTTGCAGCACTGGTCCACCA 718
Db 223 ProAlaProAlaArgArgThrAspArgAlaGlyGluArgLeuAlaAlaGlyLeuPro 242
Qy 717 GCTGGTGGCAGTGTCCGCTCAGGCTTCCAGCTGGGCTTGTCTACAGACACCC----- 664
Db 243 -----GlyProAlaProAla-----ProGlyAlaGlyHisArgGlnProGlyAla 258
Qy 663 CCAGGA-----ATCTGGCCAGCTGCTCCACCATCGTCACCAAGTCCCGATGCTT 610
Db 259 ProGlyGlyLysLeuArgAlaAlaGlyProProLeuAlaThrGlyPro----- 274
Qy 609 CATACTTGAGAAAAAGCACGTTCCAGTCCATGCGGTGCTCCC----- 568
Db 275 -----LeuArgArgGlnProAlaGlyAlaThrAlaArgLeuArgHisArg 289
Qy 567 -----AGAACTCCTGCAGCTGCTC----- 549
Db 290 ProGlyHisProAlaArgLeuArgArgGlnProAlaArg-LeuArgArgSerGlyGluArg 309
Qy 548 -----AAACACGAGGCGGT 535
Db 309 gProAlaGluSerProArgArgThrProGlyGluLeuArgGlyArgArgProAspAlaLe 329
Qy 534 AGCCACGCTTATCATTAACCT-----CCGCGACAATTTTGAAGGTGCCTC 484
Db 329 uProArgLeu---AlaGlyLysProGlyArgTyrProAlaGluPheAlaArgGlyAlaAl 348
Qy 483 GGTAGCTCATGTCGCGAGAGCGGTGGAAGTCACTGATAATAAGACACCAAGATCTTGG 424
Db 348 aProAlaProGlyProLeuAlaAlaAla-----ValProArgArgHis----- 362
Qy 423 GGTTCGAGCCATATATGATGACCTTGGAGTCTCCATTTGTGGAGTTCAGAGGGCAGAA--- 367
Db 423 GGTTCGAGCCATATATGATGACCTTGGAGTCTCCATTTGTGGAGTTCAGAGGGCAGAA--- 367
```

```
Db 363 ----AlaArgAlaAlaGlu--ProTrpArgArgProAlaGlyGlyArgAlaAspArg 380
Qy 366 -----ACGGTAGGGCAGGTGGTCT 346
Db 381 ThrAlaThrAlaProAlaArgProArgThrAlaCysAlaAlaAlaValAlaGlyProArg 400
Qy 345 TGATGAGCGGGAGAGGTTCAGTTCTTGATGATGTCAGGCCCG------GCT 298
Db 401 -----HisAlaArgA-gGluGlnProArgThrLeuProGlyProArgProGlyAla 418
Qy 297 GTGGGTACT-----CCAGGACCGGGAGGTGCTCGTCGATGT 262
Db 419 ArgGlyThrHisLeuAlaGlySerAlaGlnProAlaAlaProAlaAlaArgProPro 438
Qy 261 TCATCA-----AGCCGATCTCATCGGGTCAGGCCCTGGCTCACCAGTAGACCACT 208
Db 439 AlaAlaLeuProAlaArgAlaArgGlyArgArgGlnProProGlyArgProPro 458
Qy 207 -----CCTGCAGCAGC----- 196
Db 459 ArgArgLeuArgProGlyAlaProValAlaAlaAlaAlaProAlaThrGlyAlaAla 478
Qy 195 -----TCGTCCCGACTTG------GGT 178
Db 479 ThrGluProGlyAlaGlyAspSerArgHisArgCysArgSerGlyThrAspArgProGly 498
Qy 177 AGGTGACGA-----TCCACAGCTCGTGGCGCCGACCGGGAAGTTGGCGATCTCTCCA 124
Db 499 ArgProArgGlnProAlaAlaAlaArgAspProGlyProGlySerTrpProGlyProSer 518
Qy 123 TCTTCCCGC-----GGCAGAAGCGCGCCAGCCCAT 88
Db 519 AlaAlaArgArgAlaAlaArgArgGlyAspProHisArgArgGluThrAlaLeuArgSer 538
Qy 87 GGAACCTCGAAGTACTTGCTCTCGAACT---CCCCCGGGGTGCTGGGGGTCTCGGCCTCGC 31
Db 539 AlaAlaArgArgProArgArgAlaAspProProGlyArgThrProAlaGlnProGly 558
Qy 30 TCTCCGCCATCGCCGCCCGT-----CGCCGTGCG 1
Db 559 GlnThrGluGlyAlaArgArgLeuAlaArgArgArg 570
```

## RESULT 77

```
US-09-252-991A-23674
; Sequence 23674, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23674
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23674
```

```
Alignment Scores:
Pred. No.: 9.25e-08 Length: 335
Score: 171.00 MatChes: 89
Percent Similarity: 33.1% Conservative: 23
Best Local Similarity: 26.3% Mismatches: 110
Query Match: 3.9% Indels: 116
DB: 2 Gaps: 19
```

US-10-768-158-1 (1-2419) x US-09-252-991A-23674 (1-335)

```
Qy 840 CCATCTTCTGTTTATACACCAAGTCAAACTTCTCATTCATGAGACCGTGAAGATGTCTCT 781
Db 15 ProSerSerProCysArgProAlaAlaThrAlaArgSer----- 27
Qy 780 TCCACAGCCCAACTCTTCCCGGCCACCGGCGCTCAGCGTTCAGCACTGGTCCA 721
Db 28 -----GlnGlyArgArgAlaGlyAlaProArgCysArgArgGlySer 41
Qy 720 CCAGCTGGTGGCAGTGTCTCCGTCCAGGG-----CTTCCAGCTGGGCCCTTGT--- 676
Db 42 AlaGlyAlaAlaGlyGlyCysArgGlyArgAlaArgSerProAlaSerProCysAla 61
Qy 675 -----CACAGG-----ACACCCCCA-----GGATCTGG--- 652
Db 62 GlyArgHisArgArgAlaArgAlaValProAlaArgArgArgArgTrpGlyThrTrpCys 81
Qy 651 -----CCAGCTGCTCCACCATCTGTCACCAAGGTCCC 622
Db 82 AlaGlySerArgGlyAlaAlaProGlyProValAlaArgProArgSerAlaProArgPro 101
Qy 621 GATGCATGTCTTCATCTTGAAGAAACAGCTTCGAGTCCA---TCCGGTGTCTCCC--- 568
Db 102 AlaAlaArgSerGlyValProGlyArgArgCysArgProGlyCysProAlaProVal 121
Qy 567 -----AGAACTCTCTGCACGTCTCAACACGAGGCGGTAGCCAGCTTATCATTCATAA 514
Db 122 AlaArgGlyThrThrCysAlaAlaSerArgThrAlaAlaProAlaAlaArgGly--- 140
Qy 513 ACCTCCGCGCAGAATTCTTGAAGGTGCTCGGTAGC-----TCATGGTCCGCA 466
Db 141 -----ArgCysGlyGlySerArgArgGlyCysArgTyrValArg 153
Qy 465 GAGAGCGGT-----GGAACGTGATAAAGACACACACAGATCTCTGG----- 424
Db 154 ArgProGlyLeuArgGlyProValArgArgArgProValArgTrpProArgThrAla 173
Qy 423 -----GGTTCCGAGCCATATAGATGACCTTGG 397
Db 174 ArgAlaThrAlaValArgArgArgProGlyGlyCys----- 186
Qy 396 AGTCTCCATTGTGGA-----GGTCAGAGGGCAGAAAGCGGTAGGCGCAGGT 352
Db 187 AlaArgArgCysGlyTrpProAlaGlyAsnAlaArgAlaArgArgAsnArgAla 206
Qy 351 GGCTCTTGATGAGCGGGGAGAGGTCACTTCCTTGATGATGTCACGCGCCGCTGTGGGT 292
Db 207 GlyArg-----ProAlaAlaGly 212
Qy 291 ACTCCAGGACCGGAGCTGCTCGTCTGATGTTTCATCAAGCCGATCTCATCGGGGTACGGC 232
Db 213 ArgAlaLeuArgSerArgSerArgArgCysAlaValAlaAlaArgSerAlaAlaArgArg 232
Qy 231 CCTGGCTCACCAAGTAGACACCTCTCTGACAGCAAGCTGGTCCGCGACTTGGGGTAGGTGA 172
Db 233 AlaArgAlaAlaAlaArgProAlaProAlaAlaPro-----AlaTyrProAlaCys 249
Qy 171 CGATCCACAGCTCGTGGCGCGACCGGGAAGTTGGCGATCTCTCT----- 127
Db 250 ArgSerGlyArgGlyAlaSerAlaProGlyAlaSerThrAlaProGlyProCysIlePro 269
Qy 126 -----CCATCTTCCCGCGCAGAAAGGCGCGCAGCCGACCATGGAACCTCGA--- 79
Db 270 AlaValArgProValAlaArgArgAlaAspAlaGlySerAlaArgLeuProProArgArg 289
Qy 78 -----AGTACTTGTCTCTCGAACTCCCCCGGGGTGCTGGGGGTCTCGGCCTCGC 31
Db 290 ArgGlnArgGlnCysSerThrArgProThrGlyCys-----ArgProArg 304
```

## RESULT 78

```
US-09-252-991A-32035
; Sequence 32035, Application US/09252991A
; Patent No. 6551795
```



```
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32035
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32035

Alignment Scores:
Pred. No.: 1,32e-07 Length: 418
Score: 170.00 Matches: 82
Percent Similarity: 35.0% Conservative: 16
Best Local Similarity: 29.3% Mismatches: 118
Query Match: 3.9% Indels: 64
DB: 2 Gaps: 12

US-10-768-158-1 (1-2419) x US-09-252-991A-32035 (1-418)
QY 798 AGACGGTGAAGATGCTCTTCCACAGCCCACTTCCCGGCGCCAGCGGCGCTCAG 739
DB 17 ArgArgProArgSerPro-----ProArgArgGlyAlaAla 28
QY 738 CGTTGCAGCACTGGTGGCCAGCTGGTGGCAGTGTCCGTCAGGGCTTCCAGCTGGG--- 682
DB 29 ArgAlaThrAlaArgProAlaLeuArgGlyProArgHnaAspArgProProAlaGlySer 48
QY 681 -----CCTGTGCAGACAGACCCCGAGAACTCGGCAGCTGTCCACCATCG 634
DB 49 SerSerSerProLeuSerArgArgValProAla-----AlaSerAlaGlyAla 65
QY 633 TCACCAAGTCCC-----GATGCATGCTTTCATCTTGAGAAAAGCAGCTTCAGATCCA 580
DB 66 ThrProGlyProValArgAspGlyCysAlaGlySerValGlnArgProGlySerAlaArg 85
QY 579 TCGCGTGCT-----CCGAGAACTCTTCGACGT 553
DB 86 AlaGlyAlaSerSerProGlyArgSerArgValProGlyProProArgSerArgAspArg 105
QY 552 GCTCAAAACCGAGCGTAGCCAGCTTATCATTCATAAACCTCCGGC-----AGAATT 499
DB 106 ArgArgValArgSerGlyArgGluSerArgHisGlyIleAlaGlyGlySerAlaArgGly 125
QY 498 CTTGAAAGTGCCCTCGGTAGCTCATGTGTCGCA-----GAGAGCGGTGGAACGTGATAAT 445
DB 126 TrpArgHisAlaGlyArgArgCysSerAlaProAlaThrGlyArgAspGly 145
QY 444 AAGACACACCCAGCATCTCTGGGGTTGGCGAGCCATATAGATCACCTTGAGTCTCCATTGT 385
DB 146 SerArgProProHisProGlyAlaAlaAspArgProGlyAlaAlaAspArgArgCys 165
QY 384 GGAGTGCAGAGGGCAGAAAGCGGTAGGCAGGTGGCTTGTGATGAGCGGGGAGAGGTCA 325
DB 166 ProGlyLeuProAlaSerProProArgArgAlaAla----- 178
QY 324 GTTCTTGATGATGTCCAGGCCCGGCTGTGGTACTCCAGGACGGGAGCTGCTCGTCA 265
DB 179 -----ProArgProAlaThrGlyValAlaGlyProAlaAlaValAspArg 193
QY 264 TGTTCATCAACGCCGATCTCATCGGGGTCAGCGCTGGCTCA-----CCAAGTAGACCA 211
DB 194 AlaAlaAlaArgGlySerCysArg---ProAlaProGlySerSerAlaProGlyValPro 212
QY 210 CCTCTGCAGCAAGCTGGTCCCGACTTGGGGTAGGTGACGATCCACACGCTCGCTGGGCC 151
```

```
DB 213 AlaProAlaAla-----ArgArgLeuProCysSerAla 223
QY 150 GCACCGGGAAGTGGCGATCTCTCCATCTTCCCGCGGCAGA----- 109
DB 224 AlaThrThrValGlyCysGlyCysProSerGlyThrGlyArgTyrArgHisAlaProPro 243
QY 108 ---AGGGCGGCGAGCCGACGCGCATGGAAGTCTCGAAGTACTTCTCTCGAACTCCCGGGG 52
DB 244 GlyProAlaSerAlaAlaArgAlaProAlaArgAlaAlaSerGlyThrProProGly 263
RESULT 79
US-09-252-991A-27249
; Sequence 27249, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27249
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27249

Alignment Scores:
Pred. No.: 1,24e-07 Length: 308
Score: 169.50 Matches: 80
Percent Similarity: 33.3% Conservative: 11
Best Local Similarity: 29.3% Mismatches: 95
Query Match: 3.9% Indels: 87
DB: 2 Gaps: 15

US-10-768-158-1 (1-2419) x US-09-252-991A-27249 (1-308)
QY 762 CCGGCGCCAGCGGCGGCTTCAGCTTGCAGCACTGGTCCACCACTGGTGGC----- 709
DB 18 ProAlaArgGlyAlaGlyProGlyArgGlyGlySerGlyProGlyArgAlaGlyArgArg 37
QY 708 AGTCTCGCTCAGGGCTTCAGCTGGGCTTGTTCACAGGACACC----- 664
DB 38 SerIleProProAlaProAlaProArgGlyProArgLeuAlaArgProSerArgGlyArgGly 57
QY 663 -----CCAGGAATCTGGCCAGCTGTCTCCACCACTCGTCA 631
DB 58 ArgThrArgAlaProGlyArgCysCysSerGlyTrpTrpSerThrGlySerProMetAla 77
QY 630 CCAGGTCGCGATGCATGCTTCTATPACTTGAAGAAAAGCAGCTTCGAGTCCATCGGTGCT 571
DB 78 ProValArgSerAlaCysArgGlyThrSerArgProValArgSerTrpPro---GlyPro 96
QY 570 CCAGAACTCTGCAGCTGTCAAAACCGAGGCGGTAGCCAGCTTATCATTCATAAAC 511
DB 97 ProAlaGlyArgAlaGlyGlyTyrGlyArgArgArgGlyPro----- 110
QY 510 TCCGGCAGAATTCTTGAAGAGTGTCTCGGTAGCTCATGTCCGAGAGAGCGGTGGAAC 451
DB 111 -----ArgArgSerValProProThrValAlaSerValAlaGly----- 123
QY 450 GATAATAAGACACACAGCATCTCTGGGGTTGGAGCCATATAGATGACCTTGGAGTCTC 391
DB 124 -----ProSerArgSerGlyAlaGlyProAlaArg----- 133
QY 390 CATTGTGAGGTTCAGAGGGCAGAAAGCGGTAGGCGAGGTGGCTCTTGTATGAGCGGGGAG 331
```



```
QY 847 CACTTTCCCATCTTCTGTTTATACACCAAGTCATAAACTTCTCATTCATCGAGACGCTGAAG 788
Db :|||||:
302 ArgTrpProVal-ArgLysProArgAsnArgLysSerHis----- 315
QY 787 ATGTCCTTCCACAGCCCAACTCTTCCCGGCCACCGGCAGGCTCAGCGTTCAGCAC 728
Db :|||||:
316 -----AspArgAlaValGlnGlyProGlyTyrrArgAlaLysGlyArg- 330
QY 727 TGGTCCACAGCTGTGTCAGTGTCTCCAGGGTTCAGAGTGGCCCTTGT- 671
Db :|||||:
331 ---ArgArgAlaGlyGlyGluAlaArgSerGlyArgProGlyAsnProAlaGlyTyPr 349
QY 670 GACACCCCGAGATCTGCCAGCTGCTCCACCATCGTCACAGTCCC- 614
Db :|||||:
349 oValProProAla- 363
QY 613 TCTTCATCTTGAGAAAAAGCAGCTTCG- 586
Db :|||||:
343 sArgHisArgArgProAlaAlaArgGlyArgArgGlyArgGlyArgGlyArgAr 383
QY 585 ---AGTCCATGCGGTGCTCCAGAACTCTCTGACGTGCTCAAAACC- 533
Db :|||||:
383 gArgArgGlyAspGlyAlaProProGlnProAlaGlyAlaProAlaGlyArgGlnArgHi 403
QY 532 CCCAGCTTATCATTCATAAACCTCCGGCAGAACTTTGAAGGTCCCTCGGTAGCTCATG 473
Db :|||||:
403 sPro- 412
QY 472 GTCCGCAGAGCGGTGAAGTAATAAGACACCA- 428
Db :|||||:
412 nProGlyAlaGluGlyArgSerGlnAlaArgHisProProAlaLeuLeuArgArgSerPr 432
QY 427 TTGGGGTTGCAGGCCATATAGATGACCTTGGAGTCTCCATTGTGGAGTGCAGAGGCAGA 368
Db :|||||:
432 oAlaGlyCysLeuGly- 449
QY 367 AAGCGGTAGGCGAGGTGCTCTTGATGAGGCGGGAG- 331
Db :|||||:
449 uArgArgAlaAlaAspSer- 466
QY 330 ---AGTFCAGTCTCTTGATGATGTCACAGGCCGCGCTGGGTACTCCA- 286
Db :|||||:
466 pTyArgGly- 481
QY 285 -GGACCGGGAGCTGCTCGATGTTTCATCAAGCGATCTCATCGGGT- 238
Db :|||||:
481 rGlyProGlyHisAlaArgProGlnGlnProAlaArgLeuAlaAlaGlyThrGlyPr 501
QY 237 -----CAGCGCCTGGCTCACCAGTAGACACCTCTCTGAGCAAG 197
Db :|||||:
501 oArgLeuProThrTrpProAlaArgProGlyAspPro- 520
QY 196 CTGCTGCCGACTTGGGTAGGTGACGATCCACAGCTCGTGGGCC- 140
Db :|||||:
520 aGlyCysArgHisArgGlyThrGlyArgArgProAlaLeuGlyAlaValAlaProGlyLe 540
QY 139 TTGGCGA- 107
Db :|||||:
540 uArgArgThrAlaAspProProValIleProAlaAlaProAlaArgSerArgAlaLeuAl 560
QY 106 GGCGGCAGCCCGACCCATGAACTCGAAGTACTTGTCTCGAATCCC- 58
Db :|||||:
560 aGlySerAlaAlaLeuAlaGlyGlnProLeuAlaGluThrLeuProCysGlnProAr 580
QY 57 -----CCGGGTGTGGGGTCTGGGCTCGC- 31
Db :|||||:
580 gLeuProAlaAspArgArgAlaThrProGlyArgArgGlyAspArgLeuArgProGlyPr 600
QY 30 -----TCTCCGCCATCGCCGCGCTCGCGCTCGC 1
Db :|||||:
600 oGlyArgAlaAspAlaAlaProGlyAlaArgArgArgCysArg 615
```

RESULT 81

```
US-09-252-991A-30343
; Sequence 30343, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30343
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30343
```

## Alignment Scores:

Pred. No.:	2,34e-07	Length:	286
Score:	166.50	Matches:	81
Percent Similarity:	33.9%	Conservative:	17
Best Local Similarity:	28.0%	Mismatches:	104
Query Match:	3.8%	Indels:	87
DB:	2	Gaps:	16

US-10-768-158-1 (1-2419) x US-09-252-991A-30343 (1-286)

```
QY 789 AGATGTCTCTCCACAGCCCAACTCTTCCCGGCCACCGGAGGCGCTCAGCGTTGCAGC 730
Db :|||||:
40 ArgCysArgArgGlySerAlaValAlaGlyCysArgAlaAlaSerArgArgSer 59
QY 729 ACTGGTCCACAGCTGTGGCGAGCTGCTCCGTCCAGGCTTCCAGCTGGCGCTTGTCCACAGG 670
Db :|||||:
60 IleSerProThrProGlyAlaAlaAla- 70
QY 669 ACACCCCGAGAACTGGCCAGCTGCTCCACCACGTCACCCAGGTCCCGATGTCATGCTTT 610
Db :|||||:
71 SerAlaGlyGlySerTrpLysAla- 82
QY 609 CATACTTGGAGAAAAAGCACAGCTTCGAGTCCATCGGT- 553
Db :|||||:
83 -----ProCysGlySerAlaGlyIleProProSerArg 93
QY 552 GCTCAAACAGGAGCGGTAGCCAGCTTATCATTCATAAACCTCCGGCAGAAATTTTGAA 493
Db :|||||:
94 PheProAspAlaArgArgSerProGlyAlaAlaArgTrpProGlySerVal- 111
QY 492 AGGTGCCTCGGTAGCTCATGTCGCCAGAGAGCGGTGGAACTGATAATAAGACACCA 433
Db :|||||:
112 AlaAlaAlaThrProTrpProArgGlyProArgProArgProArgHisTrpArgProPro 131
QY 432 GATCCTTGGGTTCCGAG- 397
Db :|||||:
132 GlyProCysAlaAlaLysAlaIleArgGlySerProAlaProTyrrArgTrpProProGly 151
QY 396 AGTCTCCATTGTGAGGTGACAGGCGAAGAGCGGTAGGCGAGGT- 349
Db :|||||:
152 SerValAlaAlaThrSerAlaArgAlaProProGlyArgArgAlaProGlyArgProGly 171
QY 348 TCTTGATGAGCGGGAGAGGTGAGTTCCTTGATGATGTCAGGCGCGCTGTGGGTACT 289
Db :|||||:
172 SerArgThrGlyCysAlaArg- 189
QY 288 CCAGACCGGAGCTGCT- 247
Db :|||||:
190 ProGlyProAlaAlaThrValProAspTrpArgAspProAlaAlaArgCys- 208
QY 246 CATCGGGGTGAGCGCCCTCGCTCA- 220
```

```
Db      209 ArgArgArgArgAlaProGlySerSerArgGlyCysArgGlyAlaGlyCysValSer 228
Qy      219 AGTAGACCACCTCTCGAGCAAGCTGTGGCGGACTTGGGGTAGGTACCATCCACACGT 160
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      229 AlaProProProGlyAlaProArgGlnAlaArgArgGlyArg-----SerArg 245
Qy      159 CGCTGGGCGCCACCGGGAAGTGGCGATCTCTCCATCTCTCCCGCGCAGAAAGGGCGCA 100
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      246 ArgTyxProAlaProHisArg---ArgAlaProProAla----- 257
Qy      99 GCCGACGCCATGGAACCTCGAAGTACTTGTCTCGAACTCCCGC-----GGGTGCTGG 46
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      258 -----ProProArgThrAlaCysAla 264
Qy      45 GGGTCTCGGCTCGCTCTCCGCCATGC 19
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      265 GlySerArgTrpArgCysArgProCys 273

RESULT 82
US-09-252-991A-24899
; Sequence 24899, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24899
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24899

Alignment Scores:
Pred No.: 2,57e-07 Length: 335
Score: 166.50 Matches: 84
Percent Similarity: 34.2% Conservatives: 21
Best Local Similarity: 27.4% Mismatches: 137
Query Match: 3.8% Indels: 65
DB: 2 Gaps: 13

US-10-768-158-1 (1-2419) x US-09-252-991A-24899 (1-335)
Qy      840 CCATCTTCTGTTTATACCAAGTCAAACTTCTCATTCATGGAGACGGTGAAGA---TGT 784
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61 ProAlaArgAlaAlaProGlyAsnValGlyArgGlyTrpProGlyAsnArgSerCys 80
Qy      783 CTTCCACGCCCACTCTCCCGCGCCACGGCAGG----- 745
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      81 AlaAlaProGlySerProAlaAlaGlyProGlyAlaAlaGlySerCys 100
Qy      744 ---CCTCAGCGTTGACGACCTGGTCCACGAGTGGTGGCAGTGTCTCCGTCCGTCAGGGGTTCCA 688
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      101 PheProArgArgSerSerAlaProArgAlaGlySerIle----- 114
Qy      687 GCTGGGCTTGTACAGACACCCCGCAGGA-----ATCTGGCCA 649
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      115 -----ArgThrProProGlyGlyArgArgThrAlaArgAlaLeuPro 129
Qy      648 GCTGCTCCACCATCGTCCACAGGTCGATGCTCTCATCTTGCATCTTGCAGAAAAACAGCT 589
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      130 GlyArgProProCysSerSerGlyPro-----CysVal-----Arg 141
Qy      588 TCGAGTCCATGC-----GGTCTCCAGAACTCTCTGCACGTGCTCAAAACACGAGGAGC 538
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      142 AlaSerProAlaAlaSerGlySerAlaProGlyArgArgProArgAlaGlyValAlaGly 161
```

```
Qy      537 CGTAGCCCGCTTATCATTCAATAACCTCGGCAGAAATCTTGAAGGTCCTCGGTAGC 478
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      162 ArgHisProArgHis-----ProAlaAlaAlaLeuArgArgGlyProArgGlySer 178
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      477 TCATGTGTCACAGAGACGGTGAACCTGATAATAAGACACACCA----- 433
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      179 AlaProGlyCysIleProAlaProProProValArgArgProProAlaAlaAlaValAla 198
Qy      432 GATCCTTGGGGTTCGAGCCATATAGATGACCTTGGAGTCTCCATCTGTGGAGGTCAGAG 373
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      199 ArgAlaArgSerCysArgProAlaArg-----GlySerArg 210
Qy      372 GCAGAAAGCGGTAGGCGAGCTGCTCTTGTATGAGCGGGGAGAG----- 328
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      211 ArgValArgGlyArgSerGlyArgGlyArgThrGlySerArgThrProArgArgPro 230
Qy      327 -----TCAGTTCCTTGATGATGTCACAGCCCGGCTGTGGGTACTCCAGGACCGGGA 277
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      231 ThrArgProThrValProProTipCysAlaArgGlyProGlyThrAlaGly---Arg 249
Qy      276 GCTGCTGCTGATGTTTCATCAAGCCGATCTCATCGGGGTGAGCCCTGGCTCACCAGT 217
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      250 AlaArgArgProAlaArgProSerArgAlaGlyArgGly-----ProGlyAlaProPro 267
Qy      216 AGACCCACTCTCTGCGACCAAGCTGTCGCCGACTTGGGGTAGGTGACGATCCACACGTGCG 157
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      268 ArgProValProAlaProLeuProGlyArgProAlaArgTyrglyAlaAlaTyrgly 287
Qy      156 TGGCGCCGACCGGGAAGTTGGCGATCTCTCCATCTTCCCGCGGACAGAAAGGCGGACGCC 97
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      288 SerSerProAlaAlaHisArgValArgProSerCysPheAlaAlaSerGlyArg 307
Qy      96 GCAGCCCATGGAACCTCGAAGTACTGCTCTCGAACTCCCGCGGCTGCTGGGGTCTCGG 37
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      308 SerArgGlyLeuArgArgGlySerProArgArgSerAlaGlyAlaAlaArgArgHis 327
Qy      36 CCTGCTCTCCGCCATGCCGC 16
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      328 ProAlaArgProProArgArg 334

RESULT 83
US-09-252-991A-23991
; Sequence 23991, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23991
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23991

Alignment Scores:
Pred. No.: 2,39e-07 Length: 244
Score: 166.00 Matches: 79
Percent Similarity: 35.8% Conservatives: 19
Best Local Similarity: 28.8% Mismatches: 85
Query Match: 3.8% Indels: 91
DB: 2 Gaps: 16

US-10-768-158-1 (1-2419) x US-09-252-991A-23991 (1-244)
```

```
QY 733 CAGACTGGTCCACGAGTGTGGGAGTGTCTCCGTCAGGGCTTCAGGTGGCGCTTGTC 674
Db 5 ArgArgTrpAlaSerSer-GlySerAlaSerAlaAlaArgArgProGlyArgLeuAr 24
QY 673 CAGGACACCCAGGAATCTGGCAGCTGCTCCACCATCGTCACCAAGTCCCGATGCATG 614
Db 24 gArg-----ProGlyProProAspArgProGly----- 34
QY 613 TCTTCATCTTGAGAAAAGCAGTTCGAGTCCATCGCGTGTCCAGAACTCTCTGCACG 554
Db 35 -----LysSerArgArgValValCys----- 41
QY 553 TGCTCAAAACAGGACCGCTAGCCAGCTTATCATTAACCTCCGCGAGAACTTCTTGA 494
Db 42 -----ProGlyTyAlaArgArgThrSerAlaArg----- 51
QY 493 AAGGTGCTCGTAGTCTCATGTCGCGCAGAGCGGTGGAATGATAATAAGACACACC 434
Db 52 ----- 52
QY 433 AGATCCTTTGGGTTGCGAGCCATATAGATGACCTTTGGAGTCTCCATTGTGGAGGTACAG 374
Db 52 oAspProGlyGlyCysSerPro---ArgGlyProArgProAlaArgArgThrGlyArgSe 71
QY 373 GGCAGAAAGCGGTAGGCGAGGTGC-----TCTTGATGAGCGCGGGAGAGGTCAGT 323
Db 71 rAlaThrSerAlaSerSerAsnGlyArgCysProAlaArgLysAlaGlyProHisSerAl 91
QY 322 TCTTTGATGATGTCAGCGCCGCTGTGGTACT-----CCAGACCGGAGCTGCTCG 269
Db 91 aArgGlyAsnSerProAlaProAlaAlaGlyArgAspTrpProArgProGlySerGlyAr 111
QY 268 TCGATGT---TCATCAAGCCGATCTCATCGGGTCAGCGGCTCGCTCCACCAAGTAGACC 212
Db 111 gGlyCysProGlySerThrAspArgHisAlaGlyAlaArgProGly-----CysArgPr 129
QY 211 ACCTCTCCAGACAGCTGGTCGGAGCTGGGTAGGTAGGTAGGTCCACACATCGCTGGGC 152
Db 129 oAlaProSerAlaArgCysGlyArgArgTyArgArgGlyLeuSer-----GlySe 146
QY 151 CGCACCGGGA---AGTTGGCATCTCTCTCCA-----TCTTCCGCGCGCAGAGG-- 106
Db 146 rGlyProGlyArgGlyThrArgArgProProHisArgAlaSerSerArgProArgArgSe 166
QY 105 -----GCGCGACCGCGCAGCCATGGAATCGA---AGTAC 74
Db 166 rLeuAlaHisCysProValArgAlaAlaProAlaSerArgArgGlySerArgProAlaAs 186
QY 73 TTGC-----TCTCGAACTCCCGGGGTGCTGGGGTCTCGGCTCGCTCT----- 28
Db 186 pCysArgTyArgGlyThrProArgProAlaTrpArgArgAlaProGlyAlaSerSerAr 206
QY 27 -----CCGCCATGCGCGCGCT 10
Db 206 gArgAlaAlaAlaGlyCysProProProThrProArgArg 219
```

## RESULT 84

```
US-09-252-991A-26942
; Sequence 26942, Application US/09252991A
; Patent No., 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26942
```

```
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26942
Alignment Scores:
Pred. No.: 3,37e-07 Length: 437
Score: 166.00 Matches: 95
Percent Similarity: 35.0% Conservative: 112
Best Local Similarity: 29.7% Mismatches: 17
Query Match: 3.8% Indels: 96
DB: 2 Gaps: 20
```

US-10-768-158-1 (1-2419) x US-09-252-991A-26942 (1-437)

```
QY 789 AGATGTCCTTCCACAGCCCAACTCTTCCCGCGCCACCGGCGAGG-----CCTCAG 739
Db 6 ArgProProProValAlaGlnLeu---ProAlaGlyArgThrGlyLysProValProArg 24
QY 738 CGTTGCACACTGTGTCCACAGCTGTGTGGCAGTGTCTCCGTCAAGGCTTCCAGTGGGCT 679
Db 25 ArgLysArgThrGlyLysProGlyAlaGlyAlaGluAspProGluGlnArgArgLysPro 44
QY 678 TGTCAAGGACACCCCGAGGAATCTGCCAGCTCTCCACCATCGTCACAGGTCCCGAT 619
Db 45 AlaArgArgLeuProProSerArgGlyAsnValGlnArg----- 57
QY 618 GCATGTCCTTCATCTTCAGAAAAGACAGTTCGAGTCCATCGCGTCT-----CCC 568
Db 58 -----GluArgCysArgGlnArgProPheArgAlaGlnCysGlnPro 71
QY 567 AGAACTCTCTCAGCTGTCAAACCCAGGAGCGGTAGCCAGCTTATCATTTATAACCTCC 508
Db 72 GlnProProValLeuValGln-----ArgProAlaArgArgSerArgSer 87
QY 507 GGCAGA---ATTCTTGAAGGTGCTCGTAGCTCAGTCCGCGAGAGAGCGGT---GGA 454
Db 88 GlyGlnSerValValGlnArg-----AlaAspProGlyArgGly 100
QY 453 ACTGATAATAAGACACCCAGCATCTCTTGGGTGTCGAGCCATATAGATGACCTTGGAGT 394
Db 101 GlyAspArgThr-----GluProLeuArgArgPro----- 111
QY 393 CTCCATTGTGGAGTTCAGAGGCGAGAAAGCGTAGGCGAGGTGGCTCTTGTAGGCGGG 334
Db 112 LeuArgCysArgSerGlnSerProAlaAlaGlyArgGlyAla-----ProGly 127
QY 333 GAGAGTCACTTCTCT---TGATGATGTCAGGCGCGGCTGTGGTACTCCAGGACCGGGA 277
Db 128 ArgHisLeuLeuProGlyAlaGlnAlaProGlyProAlaArgArgThrGlyArgProAla 147
QY 276 GCTGCT-----CGTCGATGTTTCATCAAGCGCATCTCATCG-----GGTCAGCGCCT 229
Db 148 AlaAlaProValArgArgAlaGlyGlyAlaGluArgArgArgSerGlyArgThrPro 167
QY 228 GGC-----TCACCAAGTAGACCA-----CCT 208
Db 168 GlyArgArgAlaGlyArgLeuArgProAlaLeuProGlyLysSerArgArgThrGluPro 187
QY 207 CTGTCAGCAAGCTGGTCCCGACTTGGGGTAGGTGACGATCCACACGCTCGCTGGCGCGCA 148
Db 188 ProAlaProGlyArgGlyArgThrSerCysArgGlnArgGlyAlaGlyAsn---AlaAla 206
QY 147 CCGGGAAGTTGGCGATCT-----CCTCATCTTCCCGCGCGCAGAGGCG 103
Db 207 ProGlyHisAlaArgGlyAlaLeuArgHisArgProAlaSerArgGlnArgArgAla 226
QY 102 GCAGCCGCGCCCATGGA-----ACTCGAAGTACTTGC 70
Db 227 AlaGlyGlnArgThrGlyLysArgGlyValProGlyGlnArgGlnProArgAlaAlaAla 246
QY 69 TCTCGAACTCCCGGGGTCTGGGGTCTCGGCTCGCTCT----- 28
Db 69 TCTCGAACTCCCGGGGTCTGGGGTCTCGGCTCGCTCT----- 28
```

```

Db      247 GluArgAlaGlyGlnGlyArgLeuGlyArgArgThrValAlaAaspGlnProGlyAaspPro 266
Qy      27  -----CCGCATGCCCCCGCGCGTCGCGCGCGC 1
Db      267 GlyAaspHisGlyThrAaspLeuAlaHisProHisProGlyLeuArgAsgSerArgArg 286

RESULT 85
US-09-252-991A-25699
; Sequence 25699, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25699
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25699

Alignment Scores:
Pred. No.: 3,06e-07 Length: 306
Score: 165.50 Matches: 87
Percent Similarity: 36.9% Conservative: 20
Best Local Similarity: 30.0% Mismatches: 109
Query Match: 3.8% Indels: 74
DB: 2 Gaps: 18

US-10-768-158-1 (1-2419) x US-09-252-991A-25699 (1-306)
Qy      790 AAGATGTCCTTCCACAGCCCACTTCTCCCGGCCACG-----GGCAGGGCC 743
Db      62 ArgArgSerArgGProProArgPheProArgProAlaAlaLeuArgTTrpProGlyArgAla 81
Qy      742 TCAGGTTGCAGCACTGCTCCACAGCTGGTGGCAGTGCTCGTCAGGCTTCAGCTGG 683
Db      82 ProAla---ArgHis-AlaProAlaAlaLeuArgArgAlaGlyGlySerProGlyGl 100
Qy      682 GCCTGTGTACAGGACACCCCGAGGA---ATCTGGCCAGCTGCTCCACCATCGTCACCCAGG 626
Db      100 YArgSerHisGlyTrpSerValGlyArgArgTTrpProCysArgArgArgSerArgProAr 120
Qy      625 TCCCGATGCATGCTTTCATCTTGAGAAAAAGCACAGTTCGAGTCATCGCGTGTCTCCAG 566
Db      120 gAlaGly-----ArgAlaArgGlyArgAenAlaGlyCysProAr 133
Qy      565 AACTCCTGCAGTGTCAAACACAGAGCCGTAGCCAGCTTATCATTCATATAAACCCTCCGG 506
Db      133 gArg---ArgArgGlyArgSerArgSerProArgGlyGlySerValProAlaGl 152
Qy      505 CAGATTCTTGAAGGTGCTCGGTAGCTCATGTGTCGCA-----GAGAGCGGTGGAAC 452
Db      152 YArg-----ArgTrpSerGlyAlaArgArgArgProGlyTh 164
Qy      451 T-----GATAATAAGACACCAACAGATCCTTGGGGTTGCGAGCCATATAGATGACCTTG 398
Db      164 rArgProArgProArGArgProProLeuProGlyArgThrAlaProAenAlaGlyCysArgTr 184
Qy      397 GAGTCTC-----CATGTGAGGTGAGAGGGCAGAAAGCGG 362
Db      184 pCysValAaspArgSerArgThrSerArgArgPheProAlaArgArgCysArgAlaGl 204
Qy      361 TAGGCGAGGTGGCTCTTATGATGAGCGGAGAGGTGACGTCTCTTATGATGTCCAGGCCC 302
Db      204 YArgArgProValAla-----GlyGlyProArgArgProPro-----ProGlySe 219

```

```

Qy      301 GGCTGTGGGTACTCCAGGACCGGAGCTGCTCGATGTTTCATCAAGCCGATCTCATCG 242
Db      219 rAla-----ProAlaAlaArgArgValArgCysArgThrCysAla----- 234
Qy      241 GGCTCAGCGCCTGGGTCCACCAAGTAGACACACCTCTCGCAGCAAGCTGGTGGCGGACTTG 182
Db      235 -----ArgProGlyGlySerArgArgArgAlaGlyGlyTyrfGlyAlaArgTr 252
Qy      181 G-----GGTAGGTGACGATCCACATCCAGCTGCTGGGCGCGCAC 146
Db      252 pArgTyfAlaCysArgSerArgProGlyArgCysArgAenAlaArgArg-----Pr 269
Qy      145 GGGAGTTGGCGA-----TCTCTCCATCTTCCCGCGGAGAGGGCGGCGC 98
Db      269 oGlyArgArgArgThrGlyAaspLeuSerAlaProAlaSer-----ArgSerProAl 286
Qy      97 CGCAGCGCATGGAACTCGAAGTACTTGTCTCGAACTCCCGCGGGTGTCTGGGGTCTCG 38
Db      286 aAlaLeuArgGlyArgArgAenGlySerAlaGluSerAlaProGly----- 301
Qy      37 GCCTGCTCTCCGCAATGCGCGCGCGGT 10
Db      302 -----AsnCysArgArgArg 306

RESULT 86
US-09-252-991A-26078
; Sequence 26078, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26078
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26078

Alignment Scores:
Pred. No.: 4.14e-07 Length: 511
Score: 165.50 Matches: 83
Percent Similarity: 35.2% Conservative: 18
Best Local Similarity: 28.9% Mismatches: 117
Query Match: 3.8% Indels: 69
DB: 2 Gaps: 15

US-10-768-158-1 (1-2419) x US-09-252-991A-26078 (1-511)
Qy      720 CCAGCTGTGGCAGTGTCCGTCAGGCTTCCAGCTGGCTTGTCTCAGGACACCCCA 661
Db      9 ProAlaGlyAlaLeuAla-----ArgAlaAlaProArgAlaSerValProPro 24
Qy      660 GGAATCTGGCCAGCTGTCTCA-----CCATCGTCACCAAGTCCC-----GATGCATGT 613
Db      25 -----ThrProArgAlaProArgSerProProSerProGlyProAlaArgArgProCys 42
Qy      612 CTTCATACTTGGAGAAAAAGCACAGTTCGAGT-----CCATCGGTGCTCCAGAACT 562
Db      43 CysAlaValAlaArgProAlaArgAlaGlyArgHisArgProCysArgSerProArgHis 62
Qy      561 CCTGCAGTGTCAAACACAGGAGCGGTAGCCAGCTTATCATTAACCTCCGGGCGA 502
Db      63 AlaThrHisValProProGlyAlaArgAlaProAlaArgGlyGlyArgGlyArg 82

```











```
QY 381 GGTGAGCGGAGAAACGGTAGGCGAGTGGCTTCTGATGAGCGGGAGAGCTCAGTT 322
Db   :|||:|||||:|||||
109 ArgArgAlaAspArgArgThr----- 117
QY 321 CCTTGATGATGTCAGGCGCGCTGTGGTACTCCAGGACCGGAGCTGCTCGTCGATGT 262
Db   :|||:|||||:|||||
118 -----AlaArgArgSer 122
QY 261 TCATCAAGCGATCTCATCGGGTCAGCGCCTCGCTCACCAGT-----AGACCA 211
Db   :|||:|||||:|||||:|||||:|||||
123 GluArgAlaArgSerArgGlyArgArgProAlaThrArgSerAlaSerAlaArgThr 142
QY 210 CCTCTGAGCAAGCTGTCGCGACTTGGGGTAGTGCATCCACACGTCGCTGGGCC 151
Db   :|||:|||||:|||||:|||||:|||||
143 ProArgArg-----ArgThrGlySerArgArgSerSerGluAsnArg 158
QY 150 GCACCGGAAGTTGGCGATCTCCT-----CCATCTCCCGC-----GGCAGA 109
Db   :|||:|||||:|||||:|||||
159 SerArgArgSerAlaArgSerProArgArgArgGlyProArgGlyArgAlaProGlyGly 178
QY 108 AGGCGGCGACCGCACGCGCATGGAAGTCTGCTCTCGAACTCCCGCGGGTGC 49
Db   :|||:|||||:|||||:|||||
179 ArgArgSerArgGlyArgGlyArgAlaThr-----AsnProProGlyArg 195
QY 48 TGGGGTCTCGGCTCGCTCTCCGCATGCCCGCGCGCTCCCGCTGCC 1
Db   :|||:|||||:|||||:|||||
196 ThrAlaArgArgSerArgLysProArgAlaAlaArgHisProArg 211

RESULT 93
US-09-252-991A-22391
; Sequence 22391, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22391
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22391

Alignment Scores:
Pred. No.: 4.78e-07 Length: 248
Score: 163.00 Matches: 79
Percent Similarity: 36.7% Conservative: 19
Best Local Similarity: 29.6% Mismatches: 85
Query Match: 3.7% Indels: 84
DB: 2 Gaps: 16

US-10-768-158-1 (1-2419) x US-09-252-991A-22391 (1-248)

QY 717 GCTGGTGACGAGTCTCGCTCAGGCTTCCAGCTGGG----- 682
Db   :|||:|||||:|||||
35 AlaSerSerSerProAlaArgArgProAlaProGlyAlaGlyHisAlaArg 54
QY 681 -----CCTGTGCAGGACACCCAGGAAATCGGCCAGTCTGCCACCATCG 634
Db   :|||:|||||:|||||:|||||
55 ArgArgAlaIleProAlaArgAlaProProAlaAlaTrpPro--AlaProProSer 73
QY 633 -----TCACCGAGT-----CCCGATGATGTCCTTCACTTGAGAA 598
Db   :|||:|||||:|||||:|||||
74 ArgAlaAlaGlyArgAlaGlyProGlyAlaAlaProHisAlaCysAlaTrpAla----- 91
QY 597 AAGACACGTTCCAGTCCATGCGGTGCTCCAGAACTCTCGACAGCTGCTCAAAACGAGGAGC 538
```

```
Db 92 ---AlaSerAlaSerProGlyProAlaProAlaProAlaProAla---ProArgTrpArg 109
QY 537 CGTAGCCAGCTTATCATATAAACCTCCGGCAGATTCTTGAAGGTGCTCGTCCGTAGC 478
Db   :|||:|||||:|||||:|||||
110 ArgSer-----HisAlaArgThrGly-----CysProProAla 120
QY 477 TCATGTCGCCAGAGAGCGGTGGAACGTGATAAAGACACACCATGATCTCTGGGGTGC 418
Db   :|||:|||||:|||||:|||||
121 SerArgArgGlyGlnAlaProArgArgGluArgSerAlaProPro-----Ala 136
QY 417 GAGCATATAGATGACCTTGGAGTCTCATTGTGGAGTCCAGAGGCGAGAAGCGGTAGG 358
Db   :|||:|||||:|||||:|||||
137 SerProValArg-----SerGlyArg 143
QY 357 GCAGGTGCTCTTTCATGAGCGGGGAGAGTCACTTCTTTCATGATGATCCAGGCCCGCT 298
Db   :|||:|||||:|||||:|||||
144 ProValProThrProGlyHisSerAspArg-----ProTrpProGly 157
QY 297 GTGGGT-----ACTCCAGGACCGGAGCTGCTCGTCTGATGTTCA 259
Db   :|||:|||||:|||||:|||||
158 ArgGlyGlySerAlaThrArgArgProThrProGlyThrGly---AlaArgArgSerThr 176
QY 258 TCAAGCCGATCTCATCGGGTCAGCGCCTGGGTCA-----CCAAGTAGACCACTCCT 205
Db   :|||:|||||:|||||:|||||
177 SerProProIleGlnProGlyArgAsnProAlaLysArgTrpArgAsnArgProValPro 196
QY 204 GCAGCAAGCTGGTCCCGACTTGGGTAGGTGACGATCCACACGTCGCTGGCGCCACCG 145
Db   :|||:|||||:|||||:|||||
197 AlaAlaSerValAlaValLeuProGlyAlaLysArgAlaArgSerThrGluSerPro 216
QY 144 GGAAGTTGGGATCTCTCCATCTTCCCGCGGAGGCGGCGGACGCCACGCCATGGA 85
Db   :|||:|||||:|||||:|||||
217 GlySerSerArgAla-----GlyArg----- 223
QY 84 ACTCGAAGTACTGCTCTCGAACTCCCGCGGGTGTGGGGTCTCGGCTCTCGCTCTCCG 25
Db   :|||:|||||:|||||:|||||
224 -----CysCysArgSerThrProGlySerProProLysArgGlnGlyAsnAla 239
QY 24 CCATGCGCGCGCTCGCCGT 4
Db   :|||:|||||:|||||
240 ArgCysProAspArgArg 246

RESULT 94
US-09-252-991A-30843
; Sequence 30843, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30843
; LENGTH: 663
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30843

Alignment Scores:
Pred. No.: 9.56e-07 Length: 663
Score: 162.50 Matches: 172
Percent Similarity: 28.8% Conservative: 66
Best Local Similarity: 20.8% Mismatches: 301
Query Match: 3.7% Indels: 288
DB: 2 Gaps: 37
```

US-10-768-158-1 (1-2419) x US-09-252-991A-30843 (1-6663)

Qy 2131 AGTCTCAGTAAACATCCATACAGTACTAAATAATAGAAAAATATAAATGAAATTTTCACAGA 2254  
:::|||||  
Db 18 AlaAlaArgAlaSerProTrpSerThrThrArg 29  
Qy 2253 AAGCAGCTTCCTCCACAGAAACACAGGCAAGTCTGCCGAGCCACCATGCGATGCCAC 2194  
:::|||||  
Db 30 -----ProProAlaAsp-----ArgSerGlyAlaThrAlaArgValProAlaCysArg 45  
Qy 2193 GTGCCACCT-----GCCTGG----- 2179  
|||  
Db 46 ArgValProIleAArgCysAlaAsnAlaTrpProGlyArgProMetSerSerThrSerArg 65  
Qy 2178 ---CCAGCTTCTCCCATTAACAACAACAAAACACAAAGCTTCTCTACACCGTTTCAAA 2122  
|||  
Db 66 ArgProGlyTrpSerProAlaValProAlaGlyAlaCysCysTrpLysThrArgAsnVal 85  
Qy 2121 TACAGCAACGAACGGAGACCCCTTCCACAGCAGCGACCCCTCTGACCACCACTTTGGG 2062  
|||  
Db 86 AlaglyArgSerAlaThrAspAlaTrpSerSerAlaProAlaProAlaAsnCysCys 105  
Qy 2061 TGCTTTCCAGCCACATAAATGACAGACAAACCCCGTCCCAACGGGTTGGCTGCTACTG 2002  
|||  
Db 106 LeuPheProAla-----GlyProSerPro----- 113  
Qy 2001 GCCCCGACTCTCATCGCTTCAGGAAATGTGGGTGAGAAA----- 1960  
|||  
Db 114 AlaSerProAlaProAlaAlaCysAArgArgTrpProArgAlaAlaCysHisTrpProAla 133  
Qy 1959 -----TTAGGCGATGAAGCAATAAAGACCTCACCTTTACTCCCTCAG 1918  
|||  
Db 134 SerAlaTrpTrpTrpLeuAlaProAlaAlaArgCysCysTrpProAlaProAlaProAla 153  
Qy 1917 ATACACGGA-----GGGATGAGGTGGGACAGACAGCGCGCGCTGCCAAC 1870  
|||  
Db 154 SerAlaGlyAArgAlaCysCysAlaSer-ProAsnAArgArgArgGluProTrpProPr 173  
Qy 1869 TGATCCG-----AGTGAACAGGCTACACTCGCAAAATGGTCTCC----- 1827  
|||  
Db 173 oSerProTrpAlaSerArgAlaGlyProAlaSerCysGlyArgProProAlaCysSerPr 193  
Qy 1826 -----ACGGCTCAGCGCCCTCAGTGCACAGAACAACTGAATGATAAGACCGGGGA 1774  
|||  
Db 193 oValAlaThrAlaProThrAlaThrCysSer-----ProProSe 206  
Qy 1773 TTTTAGAAATTTACTATTAAAGCATCTGCTGCCAGCAATTAGAGCTATCACTTCAGGAA 1714  
|||  
Db 206 rAlaArgSerAlaTrpLysProCysAlaCysAlaValGly----- 221  
Qy 1713 ACAAGTAAAGCTGTGGAATTCAGTCAAAATTCGAGGGCAACTCGGTCTTTGGTGCC 1654  
|||  
Db 222 -----SerAlaArgSerProAlaAsnAlaTrpProAlaAlaSerAlaTrpCysPr 238  
Qy 1653 GACTCTCCACCCACTGCATGCAATGCTTCCTGACTGCTAGCTCCACCTGCAGCGCT 1594  
|||  
Db 238 oThrCysAsnSerAlaArgProSerAlaAla-----AlaAlaSerThrAlaArgSerProSer 258  
Qy 1593 CACCAAGGAGCA-----CTTCCCTCAGCATGCTCTCCGCGGCTC 1555  
|||  
Db 258 hrAenGlyArgProAlaAlaCysProThrThrProProAlaAsnAlaPro----- 275  
Qy 1554 CTCTGCATTTGGGCTTTGGAGCCGCGCTGGAAGTCTCTGGCAGAGGAGCGTCCATCT 1495  
|||  
Db 276 -----AlaSerAlaAlaAlaAsnTrp-ProTrpTrpLysAlaProSerProAla 291  
Qy 1494 GAAGACCTAGTGAAACAATTTCTGCTCATTTCTCAGGAAGGTTCTGAAGAAGTGAACCTCGCTC 1435  
|||  
Db 291 ----- 291  
Qy 1434 AGAATAAACAGTGTCCAGATTTCCCCCAACAGACCCCAACTCACTCTCTGTTCTCACAC 1375  
|||  
Db 292 -----MetProProSerThrSerAlaThrArgLeuAlaAlaCysGlyHis--- 306

Qy 1374 TATCTCAGAACCAAGTCAGAACAGAAATGGCTGTGCCTTACATAATAAACTACAAAATA 1315  
:::|||||  
Db 307 -----AspGlyAspAlaGlyArg--- 312  
Qy 1314 AAGAGTTAACCCCTCAGATCTTTCTACCATTCGGGTGGCTCGCTCTCGATTCCTCCCTGG 1255  
|||  
Db 313 -----AlaSerGlnThrProTrp 318  
Qy 1254 AAATGAACCTTTTATTTGGTTTACTGACATTTATGTAGATTTCCAGTGAAAAGCTCTATAA 1195  
:::|||||  
Db 319 ArgAspThr----- 321  
Qy 1194 AATACATAATACGGGTTGAAAAGGCAGACATTCTAGTTGCATA----- 1150  
|||  
Db 322 -----SerProCysValProAsnCysValAsn 330  
Qy 1149 TATTACAGGCTTTATCCTTACGGTCCAGGCCATTGGAACCTGGAACCTGGAGACTGTTTGT 1090  
|||  
Db 331 TrpArgArgArgIlePro-----TrpSerAlaValAlaLysThrSerPro 345  
Qy 1089 AATCAGACATGGAGAGCTGCAGTCTTAAAGGGGAGACAGCTGCTTCGGTTGGGAATCA 1030  
|||  
Db 346 TrpProArgTrpProGlyMetProAlaGlyProArgProSerCysThrAlaValAlaAla 365  
Qy 1029 TCACACTCCCTCCGCTCACGCCCTCTTCCCTTCCCGCTGTTTCACACGCTGCTCCA 970  
|||  
Db 366 TrpAlaProAlaArgGlyGlyTyAlaAlaAlaProArg----- 378  
Qy 969 GAGTTTGTCCAGCAAGGAATAATGAATGCATACAGACTTTTGGCTAGTAGACTGCTGTG 910  
|||  
Db 379 -----AsnSerCys-----SerAla 383  
Qy 909 GGTATTGTGAGCATGACAGGTTGTTGTTCTGTTATTATAAATAAAGTCAAACTGAGGT 850  
|||  
Db 384 GlyLeuArgArgArgValArgArgSer-----AlaArgAlaTrp 398  
Qy 849 CACACTTTCCTCTCTCTGTATACCAAGTCAAACTTCTCAT-----TCATGGAGA 796  
|||  
Db 399 ArgProTrpProAlaGlyArgAlaThrProAlaSerAspAlaArgArgAspSerAlaSer 418  
Qy 795 CGGTGAAGATGTCTTCCACAGCCCAACTCTTCCCGGCCACCGGCGAGCGCTCAGCGT 736  
|||  
Db 419 ArgGlyArg---ProAlaThrAlaAlaAspHisProAlaAlaTrpValSerAlaAlaArg 437  
Qy 735 TGCAGCACTGGTCCACAGCTGGTGGCAGTCTCCGTCAGGCTTCACAGCTGGCCCTGT 676  
|||  
Db 438 ArgThrSerSerAlaProIleAlaGlySerAlaProGlySerGlyThrAlaProArgCys 457  
Qy 675 CACAGGACACCCCGAGAACTTGGCCAGCTGCCACCATCGTCACCCAGTCCCGATGCA 616  
|||  
Db 458 His-----ProValArgLysAspGlyAlaGlyProAlaSerThrGly----- 471  
Qy 615 TGCTTTCATCTTGAGAAAAAGCAGCTTCGAGTCCATCGCGTGTCCCGAGAACTCTCGCA 556  
|||  
Db 472 -----ArgSerArgArgArgTrpAlaAlaAlaProAlaAlaArgAla 485  
Qy 555 CGTGCTCAACACGAGGACCGTACCCAGCTTATCATTTCAATAAACCTCCGCGAATTCCT 496  
|||  
Db 486 ArgThrGlyIleArgArgThrAlaAlaAla----- 495  
Qy 495 GAAAGGTGCTCGTAGTCTCATGTCGCGAGAGCGGTGGAACGTAGTAATAAGACACCA 436  
|||  
Db 496 ---AlaCysTyr-----ProPro 500  
Qy 435 CCAGATCCTTTGGGTTTGGCGCCATATAGATGACCTTTGGAGTCTCCATTGTTGGAGGTGAC 376  
|||  
Db 501 ProAlaPro-----GlyArg 505  
Qy 375 AGGCAGAAAGCGGTAGGCGAGGTGGCTCTTGTATGAGCGGGGAGAGGTGAGTCTCTTGA 316  
|||  
Db 506 ArgThrProAlaAlaArgArgArgSerAlaArgThrAlaProArgProAlaIleProAla 525



US-10-768-158-1 (1-2419) x US-09-252-991A-31371 (1-272)

```
Qy 648 GCTGCTCACCATCGTACACGAGTCCCGATGCATGCTTTCATCTTGAGAAAAAGCAGT 589
      |||||
Db 24 AlaAlaThrGlyValArgProGlyPro-----HisArgGlyAlaArgGln 39
      |||||
Qy 588 TCGAGTCCATCGC---GGTGCTCCAGAACTCCTGCAGTGTCTCAACAGGCGGTAGC 532
      |||||
Db 40 SerAlaGlyCysProGlyAlaProArgProGlyArg----- 52
      |||||
Qy 531 CCAGCTTATCATTAACCTCCGGCAGAAATCTTGAAGGTCCTCGGTAGTCTATGG 472
      |||||
Db 53 ---Ala-AlaIleTrpArgProGlyAlaArgGlyArgProProAlaHisArg 71
      |||||
Qy 471 TCCGAGAGCGGTGGAACTGATAATAAGACACCACCATCCTTGGGGTTGGAGCCA 412
      |||||
Db 71 gArgValArgAlaArgPheArgSerArgSerArgHisArgAlaGlyValAla-Ala- 90
      |||||
Qy 411 TATAGATGACCTTGGAGTCTCCATGTGTGGAGTCAGAGGCGAGAAAGCGGTAGGCAAGT 352
      |||||
Db 91 -----GlyArgCysArg-----ArgL 96
      |||||
Qy 351 GGCTCTTGATGAGCGGGAGAGGTGACGTCTCTTGATGATGTCCAGGCGCGCTGTGGT 292
      |||||
Db 96 euAlaGlnLeuArgProGlyProValAspAlaAlaLeuAlaArgProGlyCys-Arg 115
      |||||
Qy 291 ACTCAGAGCCGGAGCTGCTGCTGATGTTTCATCAAGCCGATCTCATCGGGGTACGGC 232
      |||||
Db 116 ArgArgGlyProAlaAlaSerArgGlyCysAlaGlyAlaAlaGly-----Leu 131
      |||||
Qy 231 CCTGGCTCACCAAGT-----AGACCACCT-----CCTGCAGCA 199
      |||||
Db 132 ProGlyGlyProSerAlaProCysArgAlaProGlnArgProProGlyArgProAlaLeu 151
      |||||
Qy 198 AGCTGGTCCCGACTTGGGGTAGTGACGATCCACACGTCGCTGG----- 154
      |||||
Db 152 ProValAlaArgProArgProArgArgSerArgArgThrGlyProValArgArg 171
      |||||
Qy 154 ----- 154
      |||||
Db 172 ThrAlaGlyGlyThrAlaArgArgArgProGlyAlaSerProAsnProGlyAla 191
      |||||
Qy 153 -----GCCGACCGGGAAGTTGGCGATCTCTCCATCTTCCCGC 115
      |||||
Db 192 GlyValArgProArgProAlaAlaGlyArgGlySerArgArgLeuProArgHisAlaArg 211
      |||||
Qy 114 GGC-----AGAAGGGCGCAGCGCCATGGAACTCGAAGTACTTGCTCTCGAACT 61
      |||||
Db 212 GlyArgValArgArgArgAlaAlaProGlyThrArgSerHisArgThrAlaGly 231
      |||||
Qy 60 CCCCCTGGGTCTGGGGTCTCGG-----CCTCGCTCCGCCATGCGGCGC 13
      |||||
Db 232 ArgProAlaCysArgSerArgGlyLeuLeuLeuProArgLeuProProAlaArg 251
      |||||
Qy 12 CGTCCCGCTCGC 1
      |||||
Db 252 AlaArgValArg 255
```

## RESULT 97

US-09-252-991A-28132  
; Sequence 28132, Application US/09252991A  
; Patent No. 6551795

## GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 28132

; LENGTH: 333

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-28132

## Alignment Scores:

Pred. No.:	1e-06	Length:	333
Score:	160.50	Matches:	64
Percent Similarity:	35.5%	Conservative:	7
Best Local Similarity:	32.0%	Mismatches:	54
Query Match:	3.6%	Indels:	75
DB:	2	Gaps:	11

US-10-768-158-1 (1-2419) x US-09-252-991A-28132 (1-333)

```
Qy 438 CCACGATCCTTGG-----GGTTGCGAGCCATATAGATGACCTTGG 397
      |||||
Db 2 ProProAlaProArgProAlaAlaArgArgProGlyCysAlaGlyArgArg----- 18
      |||||
Qy 396 AGTCTCCATTGTGAGGTACAGGGCAGAAAGCGGTAGGCGAGTGTGCTCTTGATAGGC 337
      |||||
Db 19 -----GlyArgArgSerArgProGlyArgArgGlyArgCysSerAla 33
      |||||
Qy 336 GGGGAGAGGTTCAGTCTCTTGTATGATGTCAGGC-----CCGGCTGGGTACTCCA 286
      |||||
Db 34 GlySerArg-----ProGlyTrpArgTrpProAlaAlaThrAla 47
      |||||
Qy 285 GGACCGGAGTCTGCTCGATGATGTTATCA----- 256
      |||||
Db 48 GlyArgCysAlaGlyArgArgCysGlyArgSerArgThrAlaProSerSerIleGly 67
      |||||
Qy 255 -----AGCCGATCTCATCGGGGTGAG 235
      |||||
Db 68 GlyTrpArgHisAlaGlyGlyAspGlyArgArgArgGlyArgGlyHisArgAspAsn 87
      |||||
Qy 234 -----CGCCCTGGCTCACCA----- 220
      |||||
Db 88 ProCysArgCysSerArgAspTrpArgArgAspGlyArgProGlyAlaAlaGlyArgArg 107
      |||||
Qy 219 -----AGTAGACCACTCTCGACCAAGCTGTGTCGG-----ACTTGGGGT 178
      |||||
Db 108 ThrCysArgSerArgProGlyProProTrpHisTrpProArgArgGlySerValGly 127
      |||||
Qy 177 AGGTGAGATCCACATCGTCTGGCGCCGACCGGGAAGTTGGCGATCTCTCCATCTTCC 118
      |||||
Db 128 ArgGlyArgGlySerAlaArgTrpLeuProProGlySerAlaArgSerValAlaAlaArg 147
      |||||
Qy 117 CGCGGACAGAGGGCGGCGGCGCACCGCATCGAACTCGAAGTACTTCTCTCGAACTCCC 58
      |||||
Db 148 ArgGly---ArgAlaAsnArgCysArgArgGly-----ArgArgPro 160
      |||||
Qy 57 CCGGGGTGCTGG---GGGTCTCGGCTCTGCTCTCGCATCGCCGCGCGCTCGCGCTCGC 1
      |||||
Db 161 ProGlyArgArgProGlyArgArgCysArgThrArgProAlaAlaArgArgArgArg 180
      |||||
```

## RESULT 98

US-09-252-991A-30843

; Sequence 30843, Application US/09252991A  
; Patent No. 6551795

## GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142

```
; SEQ ID NO 30843
; LENGTH: 663
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30843

Alignment Scores:
Pred. No.:      1,51e-06      Length:      663
Score:          160.50       Matches:    93
Percent Similarity: 33.7%     Conservative: 27
Best Local Similarity: 26.1%   Mismatches: 111
Query Match:     3.6%        Indels:    125
DB:              2           Gaps:      22

US-10-768-158-1 (1-2419) x US-09-252-991A-30843 (1-663)

QY 1 GCACGCGCAGCGGGCGCATGCCGAGA-----GCCAGGCGC 39
Dbbbbb| | | | | | | | | | | | | | | | | | | | |
Db AlaThrAlaArgValProAlaCysArgArgValProIleArgCysAlaAsnAlaTrpPro 56
QY 40 AGACCCCAAGCACCCCAGGTTCGAGAGAACTTTCGAGTTCCATGGCTGCGGC 99
Dbbbbb| | | | | | | | | | | | | | | | | | | | |
Db GlyArgProMetSerSerThrArgArgProGlyTrpSerProAlaValProAlaGly 76
QY 100 TGC-----CCCCCTTCGCGCGCGGAAGA-----TGAGAGAGA 132
Dbbbbb| | | | | | | | | | | | | | | | | | | | |
Db AlaCysCysTrpLysThrArgAsnValAlaGlyArgSerAlaThrAspAlaTrpSerSer 96
QY 133 TCGCCAACTTCCCGGTGCGGCCGCCAGCACGTGT-----GGATCGTCA 174
Dbbbbb| | | | | | | | | | | | | | | | | | | | |
Db AlaPro-----AlaProAlaAsnCysCysLeuPheProAlaGlyProSer 112
QY 175 CCTACCACAAGTCCGGCACCAGCTTGTGCGAGGAGTGGTCTACT-----TGG 222
Dbbbbb| | | | | | | | | | | | | | | | | | | | |
Db ProAla--SerProAlaProAlaCysArgArgTrpProArgAlaAlaCysHisTrp 131
QY 223 TGAGCCAGGCG-----CTGACCCCATGAGATCGGCTTGATGAACATCGACGAGC 273
Dbbbbb| | | | | | | | | | | | | | | | | | | | |
Db ProAlaSerAlaTrpTrpTrpLeuAlaProAlaArgCysCys-----145
QY 274 AGCTCCCGGTCTCG-----AGTACCACACAGCGCGCTCGACATCA 315
Dbbbbb| | | | | | | | | | | | | | | | | | | | |
Db TrpProAlaProProAlaProAlaSerAlaGlyArgAlaCysCysAla 161
QY 316 TCAAGGAAC-----TGACCTCTCCCGCCTCATCAAGA-----348
Dbbbbb| | | | | | | | | | | | | | | | | | | | |
Db SerProAsnArgArgArgGluProTrpProProSerProTyrrAlaSerArgAlaGly 181
QY 349 ---GCACCTGC-----CCTACCGCTTCTGCCTCTGACCTCCCAATGGAGACTCCA 399
Dbbbbb| | | | | | | | | | | | | | | | | | | | |
Db ProAlaSerCysGlyArgProProAlaCysSerProValAlaThr-----196
QY 400 AGGTCATCTATAGGCTCGCAACCCCAAGATCTGGTGGTCTTATTATCATGTTCCACC 459
Dbbbbb| | | | | | | | | | | | | | | | | | | | |
Db -----AlaProThr 199
QY 460 GCTCTCTCGCGACCATGAGCTACCGAGGCACCTTCAAGAAATCTGCGGAGGTTTATGA 519
Dbbbbb| | | | | | | | | | | | | | | | | | | | |
Db AlaThrCysSerProProSerAlaArgSerAlaTrpLys-----212
QY 520 ATGATAAGCTGGCTACGGTCTCGTGGTTTGAGCAGCTGCAGGAGT-----564
Dbbbbb| | | | | | | | | | | | | | | | | | | | |
Db -----ProCysAlaCysAlaLysAlaValGlySerAlaArgSerProAlaAsnAlaTrp 230
QY 565 TCTGGAGCACCGCATGAGCTCGAACGTGTCTTTTCTCAAGTATGAAGACATGCATCGGG 624
Dbbbbb| | | | | | | | | | | | | | | | | | | | |
Db ProAlaAlaSerAlaTrpCysProThrCysAsnSerAlaArgProSerAlaIleAlaSer 250
QY 625 ACCTGGTGACGTGGTAGCAGCTGCCAGATTCTCTGGGGGTCTCTGTGCAAGGCCCC 684
Dbbbbb| | | | | | | | | | | | | | | | | | | | |
Db ThrAlaAtgArgSerProSerThrAsnGlyArgProAlaAlaCysProThrThrPro 270
QY 685 AGCTGGAAGCCCTGACGGAGCACTGCCACCAGC-----TGGTGG-----723
```

```
Db 113 SerProAlaArgSerSerThrAlaProCys-----SerThrArgThrProThr 129
Qy 403 TCATCTATATGGCTCGCAACCCAGGATCTGGTGTCTTATATACAGTTCCACCGCT 462
Db 130 SerSer-----AlaThrProPro-----SerThrAlaThr 139
Qy 463 CTCTCGGACCATGAGCTACCGAGCACCTTTCAAGAATTCTGCCGAGGTTTATGAATG 522
Db 140 ValAlaArgProThrProProArgThrPro----- 149
Qy 523 ATAAAGCTGGGTACGGCTCTCGTTTGAGCACGTGAGAGTCTTGAGGACCCGCGATGG 582
Db 150 -----ProThrAlaValSerAlaGlyProArgGluTrp 160
Qy 583 ACTCGAAGTCTTTTCTCAAGTATGAACATGCATCGGAGCCTGCTGACGATGGTGG 642
Db 161 SerProSerArg-----ArgSerTrp 167
Qy 643 AGCAGCTGGCCAGATTCTCGGGGTCTCTGTGACAAGGCCAGCTGGAAGCCCTGACGG 702
Db 168 ThrProTrpProAlaAlaTrp-----AlaArgIleArgTrpArgCysAlaSer 183
Qy 703 AGCACTGCCACAGCTGTGTG-----ACCAAGTGTGCAACGCTGAGGCC 747
Db 184 SerThrThrThrAlaArgThrSerAlaThrSerProIleThrThrArgPro 200

RESULT 100
US-09-252-991A-30009
; Sequence 30009, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30009
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30009

Alignment Scores:
Pred. No.: 9,18e-07 Length: 195
Score: 159.50 Matches: 75
Percent Similarity: 36.7% Conservative: 15
Best Local Similarity: 30.6% Mismatches: 78
Query Match: 3.6% Indels: 77
DB: 2 Gaps: 15

US-10-768-158-1 (1-2419) x US-09-252-991A-30009 (1-195)
Qy 756 CCACGGCGAGCGGCTCTCAG-----CGTTGCAGCACTGGTCCACCAGCTGTGGCAGTGCT 703
Db 11 ProArgSerValProAlaAlaLeuArgCysArgAlaProProProAlaArgSerAlaGly 30
Qy 702 CCGTCAGGGCTTCCAGCTGGGCTTGTCCAGGACACCCCGCAGGATCTGGCCAGCTGCT 643
Db 31 CysSerAlaCysThrArgGly---CysSerArg-----ProGlyCysTrpProAlaAla 47
Qy 642 CCACCATCGTCACAGGTCCTCGATCA-----TGCTTTCATACCTTGAGAAAAA 595
Db 48 ArgProValAlaProGlyAsnAspAlaSerAlaProGlyCysLeuArgArgAspGlyArg 67
Qy 594 GCACGTTCCAGTCCATGCGGTGCTCCCGAGAACTCTCGACGTGCTCAAAACGAGGACCGT 535
Db 68 SerArgArg-----ArgThrGlyCysArgArgThrAlaAlaAla 81
```

```
Qy 534 AGCCACGTTTATCATTAACCTCGCGCAGAACTTCTGAAAGGTGCTCGGTAGTCA 475
Db 82 SerProProGlyGlySerArgArgTyrGlyArgGlyTyrGlyArgAlaArgGlyArgSer 101
Qy 474 TGGTCCGACAGAGCGGTGGAACCTGATAATAAGACACACCAGATCCT----- 427
Db 102 ArgArgProGlyArgAlaProArgProValArgLeuProAlaProGlyArgGlyPro 121
Qy 426 -----TGG-----GGTTGGAGCCATATAGATGACCTTGGAGTCTCCATTGGA 382
Db 122 GlySerSerArgTrpProGlyArgGluProAlaArg-----ArgSerValProSerGly 139
Qy 381 GGTCCAGAGGCGAGAAAGCGGTAGGGCAGGTGGCTCTTGTATGAGCGCGGAGAGGTCA 322
Db 140 ProArgArgArgArgArgGlyArgGlyGly----- 150
Qy 321 CCTTGATGATGTCAGGCCCGGCTGTGGGTACTCCAGACCGGAGAGCTGCTCGTGTGT 262
Db 151 -----GlySerProGly-----TyrArgSer 157
Qy 261 TCATCAAGCCGATCTCATCGGGTCAGCGGCTGCTCACCAGTAGACACCACTCTCTGCA 202
Db 158 ThrAlaSerArgAsp-----GlyArgProAlaArgLeuProProAla 171
Qy 201 GCAAGCTGGTCCGGACTTGGGGTAGGTGACGATCCACACGTCGCTGGCGCCGACCGGA 142
Db 172 -----Trp---ArgArgArgGlyArg----- 177
Qy 141 AGTTGGCGATCTCTCCATCTTCCCGCGCAGAAAGGCGGCGCAGCCGACCATGGAAC 82
Db 178 -----ProProArgSerAspGly-----LeuSerHisLeuAla 188
Qy 81 CGAAGTACTTGCTCT 67
Db 189 ThrAlaArgCysSer 193

Search completed: May 17, 2006, 10:01:45
Job time : 170.5 secs
```



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☐ BLACK BORDERS

☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

☐ FADED TEXT OR DRAWING

☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING

☐ SKEWED/SLANTED IMAGES

☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS

☐ GRAY SCALE DOCUMENTS

☐ LINES OR MARKS ON ORIGINAL DOCUMENT

☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**